

Table 1

BCA4 DNA sequence (SEQ ID NO:1)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset
Accession #: D13666; Nucleic Acid Accession #: NM_006475; Coding sequence: 12-2522 (start
and stop codons underlined)

```

AGAGACTCAA GATGATTCCC TTTTACCCA TGTTTTCTCT ACTATTGCTG CTTATTGTTA    60
ACCCATATAA CGCCAACAAT CATTATGACA AGATCTTGGC TCATAGTCGT ATCAGGGGTC    120
GGGACCAAGG CCCAAATGTC TGTGCCCTTC AACAGATTTT GGGCACCAAA AAGAAATACT    180
TCAGCACTTG TAAGAACTGG TATAAAAAGT CCATCTGTGG ACAGAAAACG ACTGTTTTAT    240
ATGAATGTTG CCCTGGTTAT ATGAGAATGG AAGGAATGAA AGGCTGCCCA GCAGTTTTCG    300
CCATTGACCA TGTTTATGGC ACTCTGGGCA TCGTGGGAGC CACCACAACG CAGCGCTATT    360
CTGACGCCTC AAAACTGAGG GAGGAGATCG AGGGAAAGGG ATCCTTCACT TACTTTGCAC    420
CGAGTAATGA GGCTTGGGAC AACTTGGATT CTGATATCCG TAGAGGTTTG GAGAGCAACG    480
TGAATGTTGA ATTACTGAAT GCTTTACATA GTCACATGAT TAATAAGAGA ATGTTGACCA    540
AGGACTTAAA AAATGGCATG ATTATTCCTT CAATGTATAA CAATTTGGGG CTTTTCATTA    600
ACCATATACC TAATGGGGTT GTCACTGTTA ATTGTGCTCG AATCATCCAT GGGAAACCAGA    660
TTGCAACAAA TGGTGTGTGC CATGTCAATT ACCGTGTGCT TACACAAATT GGTACCTCAA    720
TTCAAGACTT TATGAAGTAC GAAGATGACC TTTCATCTTT TAGAGCAGCT GCCATCACAT    780
CGGACATATT GGAGGCCCTT GGAAGAGACG GTCACCTCAC ACTCTTGTCT CCCACCAATG    840
AGGCTTTTGA GAAACTTCCA CGAGGTGTCC TAGAAAGGTT CATGGGAGAC AAAGTGGCTT    900
CCGAAGCTCT TATGAAGTAC CACATCTTAA ATACTCTCCA TTGTTCTGAG TCTATTATGG    960
GAGGAGCAGT CTTTGAGACG CTGGAAGGAA ATACAATTGA GATAGGATGT GACGGTGACA   1020
GTATAACAGT AAATGGAATC AAAATGGTGA ACAAAAAGGA TATTGTGACA AATAATGGTG   1080
TGATCCATTT GATTGATCAG GTCCTAATTC CTGATTCTGC CAAACAAGTT ATTGAGCTGG   1140
CTGGAAAACA GCAAAACCAC TTCACGGATC TTGTGGCCCA ATTAGGCTTG GCATCTGCTC   1200
TGAGGCCAGA TGGAGAATAC ACTTTGCTGG CACCTGTGAA TAATGCATTT TCTGATGATA   1260
CTCTCAGCAT GGTTCAGCGC CTCCTTAAAT TAATTCTGCA GAATCACATA TTGAAAGTAA   1320
AAGTTGGCCT TAATGAGCTT TACAACGGGC AAATACTGGA AACCATCGGA GGCAAAACAGC   1380
TCAGAGTCTT CGTATATCGT ACAGCTGTCT GCATTGAAA TTCTATGCATG GAGAAAGGGA   1440
GTAAGCAAGG GAGAAACGGT GCGATTACCA TATTCCGCGA GATCATCAAG CCAGCAGAGA   1500
AATCCCTCCA TGAAAAGTTA AAACAAGATA AGCGCTTTAG CACCTTCCTC AGCCTACTTG   1560
AAGCTGCAGA CTTGAAAGAG CTCTGACAC AACCTGGAGA CTGGACATTA TTTGTGCCAA   1620
CCAATGATGC TTTTAAGGGA ATGACTAGTG AAGAAAAAGA AATTCTGATA CGGGACAAAA   1680
ATGCTCTTCA AAACATCAT CTTTATCACC TGACACCAGG AGTTTTCATT GGGAAAAGGAT   1740
TTGAACCTGG TGTTACTAAC ATTTTAAAGA CCACACAAGG AAGCAAAATC TTTCTGAAAG   1800
AAGTAAATGA TACACTTCTG GTGAATGAAT TGAATCAAA AGAATCTGAC ATCATGACAA   1860
CAAATGGTGT AATTCACTGT GTAGATAAAC TCCTCTATCC AGCAGACACA CCTGTTGGAA   1920
ATGATCAACT GCTGGAATAA CTTAATAAAT TAATCAAATA CATCCAAATT AAGTTTGTTT   1980
GTGGTAGCAC CTTCAAAGAA ATCCCCGTGA CTGTCTATAC AACTAAAATT ATAACCAAAG   2040
TTGTGGAACC AAAAATTAAA GTGATTGAAG GCAGTCTTCA GCCTATTATC AAAACTGAAG   2100
GACCCACACT AACAAAAGTC AAAATTGAAG GTGAACCTGA ATTCAAGACTG ATTAAGAAG   2160
GTGAACCAAT AACTGAAGTG ATCCATGGAG AGCCAATTAT TAAAAAATAC ACCAAAATCA   2220
TTGATGGAGT GCCTGTGGAA ATAAGTAAA AAGAGACACG AGAAGAACGA ATCATTACAG   2280
GTCTGAAAT AAAATACACT AGGATTTCTA CTGGAGGTGG AGAAACAGAA GAAACTCTGA   2340
AGAAATTGTT ACAAGAAGAG GTCACCAAGG TCACCAAATT CATGGAAGGT GGTGATGGTC   2400
ATTTATTTGA AGATGAAGAA ATTAAGAAGC TGCTTCAGGG AGACACACCC GTGAGGAAGT   2460
TGCAAGCCAA CAAAAAGTT CAAGGTTCTA GAAGACGATT AAGGGAAGGT CGTTCTCAGT   2520
GAAAAATCAA AAACAGAGAA AAAATGTTTA TACAACCTTA AGTCAATAAC CTGACCTTAG   2580
AAAATTGTGA GAGCCAAGTT GACTTCAGGA ACTGAAACAT CAGCACAAAG AAGCAATCAT   2640
CAAATAATTC TGAACACAAA TTTAATATTT TTTTCTCTGA ATGAGAAACA TGAGGGAAT   2700
TGTTGAGTTA GCCTCCTGTG GTAAAGGAAT TGAAGAAAAT ATAACACCTT ACACCTTTT   2760
TCATCTTGAC ATTAAGAATT CTGGCTAACT TTGGAATCCA TTAGAGAAAA ATCCTTGTC   2820
CCAGATTTCAT TACAATTCAA ATCGAAGAGT TGTGAAGTGT TATCCATTG AAAAGACCGA   2880
GCCTTGATG TATGTTATGG ATACATAAAA TGCACGCAAG CCATTATCTC TCCATGGGAA   2940
GCTAAGTTAT AAAAATAGGT GCTTGGTGTA CAAAACCTTT TATATCAAAA GGCTTTGCAC   3000
ATTTCTATAT GAGTGGGTTT ACTGGTAAAT TATGTTATTT TTTACAACCTA ATTTTGTA   3060
CTCAGAATGT TTGTCATATG CTTCTTGCAA TGCATATTTT TTAATCTCAA ACGTTTCAAT   3120
AAAACCATTT TTCAGATATA AAGAGAATTA CTTCAAATTG AGTAATTCAG AAAAATCTAA   3180
GATTTAAGTT AAAAAGTGGT TTGGACTTGG GAA

```

BCA4 Protein sequence (SEQ ID NO:2)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset
Accession #: D13666; Protein Accession #: NP_006466; Predicted Signal sequence: 1-21; TM
domains: none; PFAM domains: fasciclin_domains: 94-232, 234-367, 496-630; Summary: a
secreted protein involved in adhesion and osteoblast development; may participate in
preferential metastasis of breast cancer to bone.

```

MIPFLPMFSL LLLLIVNPIN ANNHDKILA HSRIRGRDQG PNVCALQQIL GTKKKYFSTC    60
KNWYKKSICG KFTTVLYECC PGYMRMEGMK GCPAVLPIDH VYGTGLIVGA TTTQRYSDAS    120
KLREIEGKG SFTYFAPSNE AWDNLDSDIR RGLSENVNVE LLNALHSHMI NKRMLTKDLK    180
NGMIIPSMYN NLGLFINHYP NGVVTVNCAR IIHGNQIATN GVHVHIDRVL TQIGTSIQDF    240
IEAEDDLSSF RAAAITSDIL EALGRDGHFT LFAPTNEAFE KLPRGVLERF MGDKVASEAL    300
MKYHILNTLQ QSESIMGAV FETLEGNTIE IGCDDGSITV NGIKMVNKKD IVTNNGVIHL    360
IDQVLIPDSA KQVIELAGKQ QTTFTDLVAQ LGLASALRPD GEYTLAPVN NAFSDDTLMS    420

```

VQRLLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKGSKQG 480
 RNGAIHIFRE IIKPAEKSLH EKLKQDKRFS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA 540
 FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKGFEPG VTNIKLTQGG SKIFLKEVND 600
 TLLVNELKSK ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST 660
 FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIIKTEGPTL TKVKIEGEPE FRLIKEGETI 720
 TEVIHGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL 780
 QEEVTKVTKF IEGGDGHLFE DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ

BCA7 DNA sequence (SEQ ID NO:3)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset
 Accession #: Z29083; Nucleic Acid Accession #: NM_006670; Coding sequence: 85-1347 (start
 and stop codons underlined)

CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCAGCCC 60
 AGCTCCGGGG AAACGCGGAG CGCGATGCCT GGGGGGTGCT CCCGGGGCCC CGCCGCGGGG 120
 GACGGGGCTC TCGGGCTGGC GCGACTAGCG CTGGTACTCC TGGGTGGGT CTCTCTGTCT 180
 TCTCCACCTT CCTCGGCATC CTCCTTCTCC TCCTCGGCGC CGTTCCTGGC TTCCGCGCTG 240
 TCCGCCAGC CCCCGCTGCC GGACCACTGC CCCGCGCTGT GCGAGTGCTC CGAGGCAGCG 300
 CGCACAGTCA AGTCGCTTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC 360
 GTGCGCAACC TCTTCTTAC CGGCAACCAG CTGGCCGTGC TCCCTGCCGG CGCCTTCGCC 420
 CGCCGGCCGC CGCTGGCGGA GCTGGCCGCG CTCAACCTCA GCGGCAGCCG CCTGGACGAG 480
 GTGCGCGCGG GCGCCTTCCA GCATCTGCCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC 540
 CCACTGGCCG ACCTCAGTCC CTTCGCTTTC TCGGGCAGCA ATGCCAGCGT CTCGGCCCCC 600
 AGTCCCCTTG TGGAACTGAT CCTGAACCAC ATCGTGCCCC CTGAAGATGA GCGGCAGAAC 660
 CGGAGCTTCG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCCGTGCACT GCAGGGGCTC 720
 CGCCGCTTGG AGCTGGCCAC CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCAA 780
 CTGCCCAGCC TCAGGCACCT GGACTTAAGT AATAATTCGC TGGTGAGCCT GACCTACGTG 840
 TCCTTCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900
 CTTCAAAATG GCACCCCTGGC TGAGTTGCAA GGTCTACCCC ACATTAGGGT TTTCTCTGGAC 960
 AACAAATCCCT GGGTCTGCGA CTGCCACATG GCAGACATGG TGACCTGGCT CAAGGAAACA 1020
 GAGGTAGTGC AGGGCAAAGA CCGGCTCACC TGTGCATATC CGGAAAAAAT GAGGAATCGG 1080
 GTCCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG 1140
 CAAACCTCTT ATGTCTTCTT TTAGCCCTGA TAGGCGCTAT TTTCTCTCTG 1200
 GTTTTGATAT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCCTGC 1260
 AGGGATCACA TGGAAAGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA 1320
 AACCTCAGTT CTAATCTCGGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACTCTG 1380
 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA 1440
 TAGATACAAC GGACTTTGAC TAAAAGCAGT GAAGGGGATT TGCTTCTCTG TTATGTAAAG 1500
 TTTCTCGGTG TGTCTGTGTA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCCTC 1560
 TTCTTTTCTT TGGAACTCCT CAACACGTAT GGAGGGATT TTCAGGTTTC AGCATGAACA 1620
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680
 ACAGATAGCA TTCAACAAAA GCTGCCTCAA CTTTTTCGAG AAAAATACTT TATTCTATAA 1740
 TATCAGTTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAAT AATTGCATCC TATAAACTGC 1800
 CTGCAGACGT TAGCAGGCTC TTCAAATAAA CTCCATGGTG CACAGGAGCA CCTGCATCCA 1860
 AGAGCATGCT TACATTTTAC TGTCTGTCAT ATTACAAAAA ATAACCTGCA ACTTCATAAC 1920
 TTCTTTGACA AAGTAAATTA CTTTTTTGAT TGCAGTTTAT ATGAAAATGT ACTGATTTTT 1980
 TTTTAAATAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAATAAAG 2040
 ATTCTTAAAA GAA

BCA7 Protein sequence (SEQ ID NO:4)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset
 Accession #: Z29083; Protein Accession #: NP_006661; Predicted Signal sequence: 1-32;
 Predicted TM domains: 357-373; PFAM domains: leucine-rich_repeats: 61-90, 119-142, 143-166,
 235-258, 259-282, 294-345;
 Summary: a type 1a TM protein of unknown function, detected in multiple cancers, with highest
 expression in breast cancer.

MPGGCSRGA AGDGRRLRLAR LALVLLGWVS SSSPTSSASS FSSAPFLAS AVSAQPPLPD 60
 QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL 120
 AALNLSGSR LDEVRAFGEH LPSLRQLDLS HNPLADLSPF AFGSGSNASVS APSPLVELIL 180
 NHIVPPEDER QNRSFEGMVV AALLAGRALQ GLRRLELASN HFLYLPDRVL AQLPSLRHLD 240
 LSNNSLVSLT YVSFRNLTHL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVF LDNNPWVCDG 300
 HMADMVTLWK ETEVVQKDR LTCAYPEKMR NRVLLELNSA DLDCDPILPP SLQTSYVFLG 360
 IVLALIGAIF LLVLYLNRKG IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSSNSDV

BCX5 DNA sequence (SEQ ID NO:5)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Nucleic Acid
 Accession #: AF160477; Coding sequence: 225-1757 (start and stop codons underlined)

GGGGAGCTCG GAGCTCCCGA TCACGGCTTC TTGGGGGTAG CTACGGCTGG GTGTGTAGAA 60
 CGGGGCCGGG GCTGGGGCTG GGTCCCCTAG TGAGACCCAA GTGCGAGAGG CAAGAAGTCT 120
 GCAGCTTCTT GCCTTCTGGG TCAGTTCCTT ATTCAAGTCT GCAGCCGGCT CCCAGGGAGA 180
 TCTCGGTGGA ACTTCAGAAA CGCTGGGCAG TCTGCCTTTC AACCATGCCC CTGTCCCTGG 240
 GAGCCGAGAT GTGGGGGCCT GAGGCCTGGC TGCTGCTGCT GCTACTGCTG GCATCATTTA 300
 CAGGCCGGTG CCCCGCGGGT GAGCTGGAGA CCTCAGACGT GGTAACTGTG GTGCTGGGCC 360
 AGGACGCAAA ACTGCCCTGC TTCTACCGAG GGGACTCCGG CGAGCAAGTG GGGCAAGTGG 420

CATGGGCTCG	GGTGGACGCG	GGCGAAGGCG	CCCAGGAAT	AGCGCTACTG	CACTCCAAAT	480
ACGGGCTTCA	TGTGAGCCCG	GCTTACGAGG	GCCGCGTGGA	GCAGCCGCGG	CCCCCACGCA	540
ACCCCTGGA	CGGCTCAGTG	CTCCTGCGCA	ACGCAGTGCA	GGCGGATGAG	GGCGAGTACG	600
AGTGCCGGGT	CAGCACCTTC	CCCGCCGGCA	GCTTCCAGGC	GCGGCTGCGG	CTCCGAGTGA	660
TGGTGCCTCC	CCTGCCCTCA	CTGAATCCTG	GTCCAGCACT	AGAAGAGGGC	CAGGGCCTGA	720
CCCTGGCAGC	CTCCTGCACA	GCTGAGGGCA	GCCCAGCCCC	CAGCGTGACC	TGGGACACGG	780
AGGTCAAAGG	CACAACGTCC	AGCCGTTCCT	TCAAGCACTC	CCGCTCTGCT	GCCGTACACT	840
CAGAGTTCCA	CTTGGTGCCT	AGCCGCAGCA	TGAATGGGCA	GCCACTGACT	TGTGTGGTGT	900
CCCATCCTGG	CCTGCTCCAG	GACCAAAGGA	TCACCCACAT	CCTCCACGTG	TCCTTCCTTG	960
CTGAGGCCTC	TGTGAGGGGC	CTTGAAGACC	AAAATCTGTG	GCACATTGGC	AGAGAAGGAG	1020
CTATGCTCAA	GTGCCCTGAGT	GAAGGGCAGC	CCCCTCCCTC	ATACAACCTG	ACACGGCTGG	1080
ATGGGCTCT	GCCCAGTGGG	GTACGAGTGG	ATGGGGACAC	TTTGGGCTTT	CCCCACTGA	1140
CCACTGAGCA	CAGCGGCATC	TACGTCTGCC	ATGTCAGCAA	TGAGTTCTCC	TCAAGGGATT	1200
CTCAGGTCAC	TGTGGATGTT	CTTGACCCCC	AGGAAGACTC	TGGGAAGCAG	GTGGACCTAG	1260
TGTCAGCCTC	GGTGGTGGTG	GTGGGTGTGA	TCGCCGCACT	CTGTCTTCTG	CTTCTGGTGG	1320
TGGTGGTGGT	GCTCATGTCC	CGATACCATC	GGCGCAAGGC	CCAGCAGATG	ACCCAGAAAT	1380
ATGAGGAGGA	GCTGACCTTG	ACCAGGGAGA	ACTCCATCCG	GAGGCTGCAT	TCCCATCACA	1440
CGGACCCGAG	GAGCCAGCCG	GAGGAGAGTG	TAGGGCTGAG	AGCCGAGGGC	CACCCTGATA	1500
GTCTCAAGGA	CAACAGTAGC	TGCTCTGTGA	TGAGTGAAGA	GCCCGAGGGC	CGCAGTTACT	1560
CCACGCTGAC	CACGGTGAGG	GAGATAGAAA	CACAGACTGA	ACTGCTGTCT	CCAGGCTCTG	1620
GGCGGGCCGA	GGAGGAGGAA	GATCAGGATG	AAGGCATCAA	ACAGGCCATG	AACCATTTTG	1680
TTCAGGAGAA	TGGGACCCTA	CGGGCCAAGC	CCACGGGCAA	TGGCATCTAC	ATCAATGGGC	1740
GGGGACACCT	GCTGTGACCC	AGGCTGTGCT	CCCTTCCCTA	GGCCTGGCTC	CTTCTGTTGA	1800
CATGGGAGAT	TTTAGCTCAT	CTTGGGGGCC	TCCTTAAACA	CCCCCATTTT	TTGCGGAAGA	1860
TGCTCCCAT	CCCACTGACT	GCTTGACCTT	TACCTCCAAC	CCTTCTGTTC	ATCGGGAGGG	1920
CTCCACCAAT	TGAGTCTCTC	CCACCATGCA	TGCAGGTCACT	TGTGTGTGTG	CATGTGTGCC	1980
TGTGTGAGTG	TTGACTGACT	GTGTGTGTGT	GGAGGGGTGA	CTGTCCGTGG	AGGGGTGACT	2040
GTGTCCGTGG	TGTGTATTAT	GCTGTATAT	CAGAGTCAAG	TGAACTGTGG	TGTATGTGCC	2100
ACGGGATTTG	AGTGGTTGCG	TGGGCAACAC	TGTCAGGGTT	TGGCGTGTGT	GTATGTGGC	2160
TGTGTGTGAC	CTCTGCCCTG	AAAAGCAGGT	ATTTTCTCAG	ACCCAGAGC	AGTATTAATG	2220
ATGCAGAGGT	TGGAGGAGAG	AGGTGGAGAC	TGTGGCTCAG	ACCCAGGTGT	GCGGGCATAG	2280
CTGGAGCTGG	AATCTGCCTC	CGGTGTGAGG	GAACCTGTCT	CCTACCACTT	CGGAGCCATG	2340
GGGGCAAGTG	TGAAGCAGCC	AGTCCCTGGG	TCAGCCAGAG	GCTTGAACCTG	TTACAGAAGC	2400
CCTCTGCCCT	CTGGTGGCCT	CTGGGCCTGC	TGCATGTACA	TATTTTCTGT	AAATATACAT	2460
GGCCCGGAG	CTTCTTGACG	GAATACTGCT	CCGAATCACT	TTTAATTTT	TTCTTTTTT	2520
TTTCTTGCCC	TTTCCATTAG	TTGTATTTT	TATTTATTTT	TATTTTAT	TTTTTTTGA	2580
GATGGAGTCT	CATCATGTTG	CTCAGGCTGG	CCTTGAACCT	CTGGGCTCAA	GCAATCCTCC	2640
TGCTCAGCC	TCCCTAGTAG	CTGGGACTTT	AAGTGTACAC	CACTGTGCCT	GCTTGAATC	2700
CTTTACGAA	AGAAAAA	AATTAAAGAA	AGCCTTTAGA	TTTATCCAAT	GTTTACTACT	2760
GGGATTGCTT	AAAGTGCCT	CCCTCCAACA	CCAGGGGGTT	AATTCCTGTG	ATTGTGAAAG	2820
GGGCTACTTC	CAAGGCATCT	TCATGCAGGC	AGCCCCCTGG	GAGGGCACCT	GAGAGCTGGT	2880
AGAGTCTGAA	ATTAGGGATG	TGAGCCTCGT	GGTTACTGAG	TAAGGTAAAA	TTGCATCCAC	2940
CATTGTTTTG	GATACCTTAG	GGAAATTGCTT	GGACCTGGTG	ACAAGGGCTC	CTGTTCAATA	3000
GTGGTGTGG	GGAGAGAGAG	AGCAGTGATT	ATAGACCGAG	AGAGTAGGAG	TTGAGGTGAG	3060
GTGAAGGAG	TGCTGGGGGT	GAGAATGTCG	CCTTCCCCC	TGGGTTTTGG	ATCACTAATT	3120
CAAGGCTCTT	CTGGATGTTT	CTCTGGGTTG	GGGCTGGAGT	TCAATGAGGT	TTATTTTATG	3180
CTGGCCAC	CAGATACACT	CAGCCAGAAT	ACCTAGATT	AGTACCCAAA	CTCTTCTTAG	3240
TCTGAAATCT	GCTGGATTTC	TGGCCTAAGG	GAGAGGCTCC	CATCCTTCGT	TCCCCAGCCA	3300
GCCTAGGACT	TGGAATGTGG	AGCCTGAAGA	TCTAAGATCC	TAACATGTAC	ATTTTATGTA	3360
AATATGTGCA	TATTTGTACA	TAAAATGATA	TTCTGTTTTT	AAATAAACAG	ACAAAACCTG	3420
TTCAAAAAA	AAAAAAA	AAAAAAA				

BCX5 Protein sequence (SEQ ID NO:6)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Protein Accession #: AF160477; Predicted Signal sequence: 1-26; Predicted TM domains: 355-371; PFAM domains: IgSF domain: 45-129, 162-225, 263-317; Summary: A type Ia TM protein; is a member of the immunoglobulin superfamily.

MPLSLGAEMW	GPEAWLLLLL	LLASFTGRCP	AGELETSADV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAVARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRVS	TFPAGSFQAR	LRLRMVPPPL	PSLNPGPALE	EGQGLTLAAS	CTAEGSPAPS	180
VTWDETVKGT	TSSRSFKHSR	SAAVTSEFHL	VPSRSMNGQP	LTCVVSHPLG	LQDQRITHIL	240
HVSFLAEASV	RGLEQNLWH	IGREGAMLC	LSEGQPPPSY	NWTRLDGPLP	SGVRVDGDTL	300
GFPPLTTEHS	GIYVCHVSN	FSSRDSQVTV	DVLDPQEDSG	KQVDLVASV	VVVGVIALL	360
FCLLVVVVVL	MSRYHRRKAQ	QMTQKYEEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
EGHPDSLKDN	SSCSVMSEEP	EGRSYSTLTT	VREIETQTEL	LSPGSGRAEE	EEDQDEGIKQ	480
AMNHVQENG	TLRAKPTGNG	IYINGRGHLV				

mouse BCX5 Protein sequence (SEQ ID NO:7)

Gene name: mouse LNIR; Unigene number: n/a; Probeset Accession #: BF168327; Protein Accession #: n/a; Predicted Signal sequence: 1-27; Predicted TM domains: 346-362; PFAM domains: IgSF_domains:44-126,166-221,259-313; Summary: This is the mouse orthologue of human BCX5; it is a type Ia TM protein of unknown function.

MPLSLGAEMW	GPEAWLRLLF	LASFTGQYSA	GELETSDVVT	VVLGQDAKL	CFYRGDPDEQ	60
QVGQVAVARV	PNEXYPGAGL	LHSKYGLHVN	PAYEDRVEQX	XHETFRRSVL	LRNAVQADEG	120

EYECRVSTFP SGSFQARMRL RVLVPLPLPSL NPGPPLEEGQ ADVAASCTAE GSPAPSVTWD 180
 TEVKGTSQSSR SFTHPRSAAV TSEFHLVPSR SMNGQPLTCV VSHPGLLQDR RITHTLQVAF 240
 LAEASVRGLE DQNLWQVGRE GATLKLSEGE QPPPKYNWTR LDGPLPSGVR VKGDTLGFPP 300
 LTTEHSGVYX CHVSNELSSR DSQVTVEVLD PEDPGKQVDL VSASVIIVGV IAALLFCLLV 360
 VVVVLSRYH RRKAQMQTKQ YEEELTLTRE NSIRRLHSHH SDPRSQPEES VGLRAEGHPD 420
 SLKDNSSCSV MSEEPEGRSY STLTTVREIE TQTELLSPGS GRTEEDDDQD EGIKQAMNHL 480
 CRKMGF

BCZ6 DNA sequence (SEQ ID NO:8)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Nucleic Acid Accession #: NM_002184;
 Coding sequence: 256-3012 (start and stop codons underlined)

GAGCAGCAA AAGCCCGCG GAGTCGCGT GGGCCGCCCG GCGCAGCTG AACCGGGGGC 60
 CGCGCCTGCC AGGCCGACGG GTCTGGCCCA GCCTGGCGCC AAGGGGTTCG TCGCTGTGG 120
 AGACGCGGAG GGTTCGAGCG GCGCGGCGCT AGTGAAACCC AATGGAAAAA GCATGACATT 180
 TAGAAGTAGA AGACTTAGCT TCAAATCCCT ACTCCTTCAC TTACTAATTT TGTGATTGG 240
 AAATATCCGC GCAAGATGTT GACGTTGCAG ACTTGGGTAG TGCAAGCCTT GTTTATTTTC 300
 CTCACCACTG AATCTACAGG TGAACCTCTA GATCCATGTG GTTATATCAG TCCTGAATCT 360
 CCAGTTGTAC AACTTCATT TAATTTCACT GCAGTTTGTG TGCTAAAGGA AAAATGTATG 420
 GATTATTTTC ATGTAAATGC TAATTACATT GTCTGGAAAA CAAACCATT TACTATTCTT 480
 AAGGAGCAAT ATACTATCAT AAACAGAACA GCATCCAGTG TCACCTTTAC AGATATAGCT 540
 TCATTAAATA TTCAGCTCAC TTGCAACATT CTTACATTGC GACAGCTTGA ACAGAATGTT 600
 TATGGAATCA CAATAATTTT AGGCTTGCCT CCAGAAAAAC CTAAAAATTT GAGTTGCATT 660
 GTGAACGAGG GGAAGAAAAA GAGGTGTGAG TGGGATGGTG GAAGGGAAAC ACACTTGGAG 720
 ACAAACTTCA CTTTAAATC TGAATGGGCA ACACACAAGT TTGCTGATTG CAAAGCAAAA 780
 CGTGACACCC CCACCTCATG CACTGTTGAT TATTCTACTG TGTATTTTGT CAACATTGAA 840
 GTCTGGGTAG AAGCAGAGAA TGCCCTTGGG AAGGTTACAT CAGATCATAT CAATTTTGAT 900
 CCTGTATATA AAGTGAAGCC CAATCCGCCA CATAATTTAT CAGTGATCAA CTCAGAGGAA 960
 CTGTCTAGTA TCTTAAATTT GACATGGACC AACCCAAAGT TTAAGAGTGT TATAATACTA 1020
 AAATATAACA TTCAATATAG GACCAAAGAT GCCTCAACTT GGAGCCAGAT TCCTCTGAA 1080
 GACACAGCAT CCACCCGATC TTCACTCACT GTCCAAGACC TTAACCTTT TACAGAATAT 1140
 GTGTTTAGGA TTCGCTGTAT GAAGGAAGAT GGTAAAGGAT ACTGGAGTGA CTGGAGTGAA 1200
 GAAGCAAGTG GGATCACCTA TGAAGATAGA CCATCTAAG CACCAAGTTT CTGGTATAAA 1260
 ATAGATCCAT CCCATACCTA AGGCTACAGA ACTGTACAAC TCGTGTGGAA GACATTGCCT 1320
 CCTTTTGAAG CCAATGGAAA AATCTTGGAT TATGAAGTGA CTCTACAAG ATGGAATCA 1380
 CATTTACAAA ATTACACAGT TAATGCCACA AAACCTGACAG TAAATCTCAC AAATGATCGC 1440
 TATCTAGCAA CCCTAACAGT AAGAAATCTT GTTGGCAAAT CAGATGCAGC TGTTTAACT 1500
 ATCCCTGCCT GTGACTTTCA AGCTACTCAC CCTGTAATGG ATCTTAAAGC ATTTCCCAAA 1560
 GATAACATGC TTTGGGTGGA ATGGACTACT CCAAGGGAAT CTGTAAAGAA ATATATACTT 1620
 GAGTGGTGTG TGTATCAGA TAAAGCACCC TGTATCAGC ACTGGCAACA AGAAGATGGT 1680
 ACCGTGCATC GCACCTATTT AAGAGGGAAC TTAGCAGAGA GCAATGCTA TTTGATAACA 1740
 GTTACTCCAG TATATGCTGA TGGACCAGGA AGCCCTGAAT CCATAAAGGC ATACCTTAAA 1800
 CAAGATCCAC CTTCCAAAGG ACCTACTGTT CGGACAAAAA AAGTAGGGAA AAACGAAGCT 1860
 GTCTTAGAGT GGGACCAACT TCCTGTTGAT GTTCAGAATG GATTATCAG AAATATACT 1920
 ATATTTTATA GAACCATCAT TGGAAATGAA ACTGCTGTGA ATGTGGATTG TTCCACACA 1980
 GAATATACAT TGCTCTCTTT GACTAGTGAC ACATTGTACA TGGTACGAAT GGCAGCATA 2040
 ACAGATGAAG GTGGGAAGGA TGGTCCAGAA TTCACTTTTA CTACCCCAAA GTTTGCTCAA 2100
 GGAGAAATG AAGCCATAGT CGTGCTGTG TTGCTTAGCAT TCCTATTGAC AACTCTTCTG 2160
 GGAGTGCTGT TCTGCTTTAA TAAGCGAGAC CTAATTAATA AACACATCTG GCCTAATGTT 2220
 CCAGATCCCT CAAAGAGTCA TATTGCCAG TGGTCACCTC ACACCTCTCC AAGGCACAAT 2280
 TTTAATTCAA AAGATCAAT GTATTAGAT GGCAATTTCA CTGATGTAAG TGTGTGGAA 2340
 ATAGAAGCAA ATGACAAAAA GCCTTTTCCA GAAGATCTGA AATCATTGGA CCTGTTCAA 2400
 AAGGAAAAAA TTAATACTGA AGGACACAGC AGTGGTATTG GGGGGTCTTC ATGCATGTC 2460
 TCTTCTAGGC CAAGCATTTT TAGCAGTGAT GAAATGAAT CTTCAAAAA CACTTCGAGC 2520
 ACTGTCCAGT ATTCTACCGT GGTACACAGT GGCTACAGAC ACCAAGTTCC GTCAGTCCAA 2580
 GTCTTCTCAA GATCCGAGTC TACCAGCCCC TTGTTAGATT CAGAGGAGCG GCCAGAAGAT 2640
 CTACAATTAG TAGATCATGT AGATGGCGGT GATGGTATTT TGCCAGGCA ACAGTACTTC 2700
 AAACAGAACT GCAGTCAGCA TGAATCCAGT CCAGATATTT CACATTTTGA AAGGTCAAAG 2760
 CAAGTTTCAT CAGTCAATGA GGAAGATTTT GTTAGACTTA AACAGCAGAT TTCAGATCAT 2820
 ATTTTCAAA CTGTGGATC TGGGCAATG AAAATGTTT AGGAAGTTTC TGCAGCAGAT 2880
 GCTTTTGGTC CAGGTACTGA GGGACAAGTA GAAAGATTTG AAACAGTTGG CATGGAGGCT 2940
 GCGACTGATG AAGGCATGCC TAAAAGTTAC TTACCACAGA CTGTACGGCA AGGCGGTAC 3000
 ATGCCTCAGT GAAGGACTAG TAGTTCTCTG TACAACCTCA GCACTACCTA TAAAGTAAAG 3060
 CTAATATGAT TTTATCTGTG AATTC

BCZ6 Protein sequence (SEQ ID NO:9)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Protein Accession #: NP_002175; Predicted
 Signal sequence: 1-22; Predicted TM domains: 625-641; PFAM domains:
 fibronectin_type_III_domains: 222-311, 424-509, 519-606; Summary: A type I TM protein; it
 homodimerizes or heterodimerizes to make a functional receptor for IL-6, oncostatin-M, IL-11,
 LIF, and CNTF.

MLTLQTVVQV ALFIFLTES TGEILLPCGY ISPESPVVQL HSNFTAVCVL KEKCMDYFHV 60
 NANYIVWKTN HFTIPKEQYT IINRTASSVT FTDIASLNIQ LTCNLTFGQ LEQNVYGITI 120

ISGLPPEKPK	NLSICVNEGK	KMRCEWDGGR	ETHLETNFTL	KSEWATHKFA	DKAKARDTPT	180
SCTVDYSTVY	FVNIEVWVEA	ENALGKVTS	HINFPDVKV	KPNPPHNLVS	INSEELSSIL	240
KLTWTNPSIK	SVIILKYNIQ	YRTKDASTWS	QIPPEDTAST	RSSFVQDLK	PFTEYVFRIR	300
CMKEDGKGYW	SDWSEEASGI	TYEDRPSKAP	SFWYKIDPSH	TQGYRTVQLV	WKTLPPEAN	360
GKILDYEVTL	TRWKSHLQNY	TVNATKLTVN	LTNDRYLATL	TVRNLVGKSD	AAVLTIACD	420
FQATHPVMDL	KAFPKDNMLW	VEWTTPRESV	KKYILEWCVL	SDKAPCITDW	QQEDGTVHRT	480
YLRGNLAESK	CYLITVTPVY	ADGPGSPESI	KAYLKQAPPS	KGPTVRTKKV	GKNEAVLEWD	540
QLPVDVQNGF	IRNYTIFYRT	IIGNETAVNV	DSSHTEYTLS	SLTSDTLYMV	RMAAYTDEGG	600
KDGPFTFTT	PKFAQGEIEA	IVVPVCLAF	LTTLLGVLCF	FNKRDLIKKH	IWPNVPDPSK	660
SHIAQWSPH	PPRHNFSK	QMYSDGNFTD	VSVVEIEAND	KKPFPEDLKS	LDLFFKKEKIN	720
TEGHSSGIGG	SSCMSSSRPS	ISSSDENESS	QNTSSSTVQYS	TVVHSGYRHQ	VPSVQVFSRS	780
ESTQPLLDSE	ERPEDLQVLD	HVDGGDGILP	RQQYFKQNC	QHSSPDISH	FERSKQVSSV	840
NEEDFVRLKQ	QISDHISQSC	GSGQMKMFQE	VSAADAFGPG	TEGQVERFET	VGMEATDEG	900
MPKSYLPQTV	RQGGYMPQ					

BFG4 DNA sequence (SEQ ID NO:10)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
Nucleic Acid Accession #: AB020689; Coding sequence: 108-2777 (start and stop codons underlined)

GAACCTATGT	AGCCTCATT	TCCCCTCCG	TGAGGTGACA	ATTGTGGAAA	AGGCAGACAG	60
CTCCAGTGTG	CTCCCCAGTC	CCTTATCACA	TCAGCACCCG	AAACAGGATG	ACCTTCCTAT	120
TTGCCCAACTT	GAAAGATAGA	GACTTCTAG	TGCAGAGGAT	CTCAGATTTC	CTGCAACAGA	180
CTACTTCCAA	AATATATTCT	GACAAGGAGT	TGTCAGGAA	TTACAACAGT	TCAGATGATG	240
AGGTGTACTC	TCGACCCAGC	AGCCTCGTCT	CCTCCAGCCC	CCAGAGAAGC	ACGAGCTCTG	300
ATGCTGATG	AGAGCGCCAG	TTTAACCTAA	ATGGCAACAG	CGTCCCCACA	GCCACACAGA	360
CCCTGATGAC	CATGTATCGG	CGGCGGTCTC	CCGAGGAGTT	CAACCCGAAA	TTGGCCAAAG	420
AGTTTCTGAA	AGAGCAAGCC	TGGAAGATT	ACTTTGCTGA	GTATGGGCAA	GGGATCTGCA	480
TGTACCGCAC	AGAGAAAACG	CGGGAGCTGG	TGTTGAAGGG	CATCCCGGAG	AGCATGCGTG	540
GGGAGCTCTG	GCTGCTGCTG	TCAGGTGCCA	TCAATGAGAA	GGCCACACAT	CCTGGGTACT	600
ATGAAGACCT	AGTGGAGAAG	TCCATGGGGA	AGTATAATCT	CGCCACGGAG	GAGATTGAGA	660
GGGATTTACA	CCGCTCCCTT	CCAGAACACC	CAGCTTTTCA	GAATGAAATG	GGCATTGCTG	720
CACTAAGGAG	AGTCTTAACA	GCTTATGCTT	TTCGAAATCC	CAACATAGGG	TATTGCCAGG	780
CCATGAATAT	TGTCACCTCA	GTGCTGCTGC	TTTATGCCAA	AGAGGAGGAA	GCTTTCTGGC	840
TGCTTGTGGC	TTTGTGTGAG	CGCATGCTCC	CAGATTACTA	CAACACCAGA	GTGTGTGGTG	900
CACTGGTGGA	CCAAGGTGTC	TTTGAGGAGC	TAGCACGAGA	CTACGTCCCA	CAGCTGTACG	960
ACTGCATGCA	AGACCTGGGC	GTGATTTC	CCATCTCCCT	GTCTTGTTTC	CTCACACTAT	1020
TTCTCAGTGT	GATGCCTTTT	GAGAGTGCAG	TGTGTGTTGT	TGACTGTTTC	TTCTATGAAG	1080
GAATTAAGT	GATATTCAG	TTGGCCCTAG	CTGTGCTGGA	TGCAAATGTG	GACAACTGT	1140
TGAACCTGCA	GGATGATGGG	GAGGCCATGA	CCGTTTTGGG	AAGGTATTTA	GACAGTGTGA	1200
CCAATAAAGA	CAGCACACTG	CCTCCCATTC	CTCACCTCCA	CTCCTTGCTC	AGCGATGATG	1260
TGGAACCTTA	CCCTGAGGTA	GACATCTTTA	GACTCATCAG	AACTTCCTAC	GAGAAATTCG	1320
GAACCTATCCG	GGCAGATTG	ATTGAACAGA	TGAGATTCAA	ACAGAGACTG	AAAGTGATCC	1380
AGACGCTGGA	GGATACCTAC	AAACGCAACG	TGGTACGAAC	CATTGTGACA	GAAACTTCCT	1440
TTACCATTGA	TGAGCTGGAA	GAACTTTATG	CTCTTTTCAA	GGCAGAACAT	CTCACCAGCT	1500
GCTACTGGGG	CGGGAGCAGC	AACGCGCTGG	ACCGGCATGA	CCCCAGCCTG	CCCTACCTGG	1560
AACAGTATCG	CATTGACTTC	GAGCAGTTCA	AGGGAATGTT	TGCTCTTCTC	TTTCTTGGG	1620
CATGTGGAAC	TCACTCTGAC	GTTCTGGCCT	CCGCTTGT	CCAGTTATTA	GATGAAAATG	1680
GAGACTCTTT	GATTAACCTC	CGGGAGTTTG	TCTCTGGGCT	AAGTGCTGCA	TGCCATGGGG	1740
ACCTCACAGA	GAAGCTCAAA	CTCCTGTACA	AAATGCACGT	CTGCTCTGAG	CCATCCTCTG	1800
ATCAAGATGA	ACCAGATTCT	GCTTTTGAAG	CAACTCAGTA	CTTCTTTGAA	GATATTACCC	1860
CAGAATGTAC	ACATGTTGTT	GGATTGGATA	GCAGAAGCAA	ACAGGGTGCA	GATGATGGCT	1920
TTGTTACGGT	GAGCCTAAAG	CCAGACAAAG	GGAAGAGAGC	AAATTCCCAA	GAAAATCGTA	1980
ATTATTTGAG	ACTGTGGACT	CCAGAAAATA	AACTTAAGTC	AAAGAATGCA	AAGGATTTAC	2040
CCAAATTAAA	TCAGGGGCAG	TTCATTGAAC	TGTGTAAGAC	AATGTATAAC	ATGTTTCAGC	2100
AAGACCCCAA	TGAGCAGGAG	CTGTACCATG	CCACGGCAGC	AGTGACCAGC	CTCCTGCTGG	2160
AGATTGGGGA	GGTCCGCAAG	TTGTTCTGTG	CCAGCCTG	AAAGGAGGGC	GGGAGCGGAG	2220
GCAGTGGGCC	GTCTGCGCAC	CAGGGCATCC	CAGGCGTGCT	CTTCCCCAAG	AAAGGGCCAG	2280
GCCAGCCTTA	CGTGGTGGAG	TCTGTTGAGC	CCCTGCCGGC	CAGCCTGGCC	CCCAGACAGC	2340
AGGAACACTC	CCTTGGAGGA	CAAAATGGAG	ACATCAAGCT	GGAGGACTCC	TCGCCCCGGG	2400
ACAACGGGGC	CTGCTCCTCC	ATGCTGATCT	CTGACGACGA	CACCAAGGAC	GACAGCTCCA	2460
TGTCCTCATA	CTCGGTGCTG	AGTGCCGGCT	CCCACGAGGA	GGACAAGCTG	CACTGCGAGG	2520
AAATCGGAGA	GGACACGGTC	CTGGTGCGGA	GCGGCCAGGG	CACGGCGGCA	CTGCCCCGGA	2580
GCACCAGCCT	GGACCGGGAG	TGGGCCATCA	CCTTCGAGCA	GTTCTCTGGC	TCCCTCTTAA	2640
CTGAGCCTGC	CCTGGTCAAG	TACTTTGACA	AGCCCGTGTG	CATGATGGCC	AGGATTACCA	2700
GTGCAAAAAA	CATCCGGATG	ATGGGCAAGC	CCCTCACCTC	GGCCAGTGAC	TATGAAATCT	2760
CGGCCATGTC	CGGCTGACAC	GGGCGCCTTC	CCGGGGGAGT	GGGAGGAGAG	GGAGGGGAGG	2820
GATTTTTTAT	GTCTTCTGT	GTTGAGTTT	TTCTTTCTTT	CTTTTAAAT	AAATATTAT	2880
TAGTACCTGG	AATTGAAGCC	TAGTGTTTTC	ATAATGTAAT	TCAATGAAAA	CTGTTGGAGA	2940
AATATTTAAA	CACCTCAATG	TAGTACATT	ACACTCTTGT	TGCGGGGAGG	GGATTTACCA	3000
GAATACAGTT	TATTTCTGTA	ATTCTAAAAA	ACAAAAAGAT	GAATCTGTCA	GTGATATGTG	3060
TGTATTATAA	CTTATTAACT	TTGCTGTTGA	GCTGTATACA	TGGTTTAAAA	AATAGTACTG	3120
TTTAAATGCTA	AGTAAGGCAG	CAGTCATTTG	TGTATTACAG	CTTTTAAAT	AAAATTAGAG	3180
CTGTAAGGAA	AATGAAAAGC	CACAAATGCA	AGACTGTTCT	TAAATGGAAG	GCATAGTCAG	3240
CGAGGGTAAA	TCCTATACCA	CTTTAGGAAG	TATTAAAAAT	ATTTTAAAGA	TTTGAATAT	3300
ATTTATAGTA	AGTCTCTAT	TCAAAATCAT	ATTCCACAGA	TGTTCCCTTT	CAAAGGGAAA	3360

ACATTTGGGG TTCTAAACAG TTATGAAAGT AAGTGATTTT TACATGATTC CAGAATAACA 3420
 CTTGTATTGA CCAATTAGAG CAGATACAG ACCAATTTTG CATTTAAGAA ATTGTTCTGA 3480
 TTATTACGT CAACTCATTA GAATTCAGTG AAAAGTAACA GTCTTTTGTC ACAGAGAATC 3540
 TGAAAGTAGC AGCAAAGACA GAGGGCTCAT GACAGGTTTT TGCTTTTGCT TTGCTTTTGT 3600
 TTTTGAAAGA GTAAAGTAC TGATGCTTCT GATACTGGAT GTTTAGCTTC TTAGTGCAA 3660
 AACATAAGTA AACAGTCAA CTTTACCATT TCCGTATTCT CCATAGATTG AAGAAATTTA 3720
 TACCACATAT CGCATATGAC CATCTTTCCA TCAAATCAAT GTAGAGATAA TGTAACCTGA 3780
 AAAAAAATCT GCAAGATAAT GTAACGAAT GTTTTAAAAA CAGAACTGT CACTTTATAT 3840
 AAAAGAATAG TATGCTCTAT TTCCTGAATG GATGTGGAAA TGAAAGCTAG CGCACCTGCA 3900
 CTTTGAATTC TTGCTTCTTT TTTATTACTG TTATGATTTT GCTTTTACA GATGTTGGAC 3960
 GATTTTCTT TCTGATTGTT GAATTCATAA TCATGGTCTC ATTTCCCTTG CTTCTTTGGA 4020
 ATATTCTTT CAACACATT CTTTATTTTA TTATACATTG TGTCCTTTT TTAGCTATTG 4080
 CTGCTGTGT TTTTATTCT TGATTTTAA ACTGTCAAAT GAAGTAGTGT 4140
 TAACCTCAA TAGGCTAAAT GTGAACAAAT AAAATACAGC AAATACTCAG AAAAAA 4200
 AAAAAA AAAAA

BFG4 Protein sequence (SEQ ID NO:11)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
 Protein Accession #: BAA74905; Signal sequence: none; Predicted TM domains: 302-318; PFAM
 domains: TBC_domain: 135-347; Summary: a Type II membrane protein, likely localized to the
 peroxisome.

MTFLFANLKD RDLVQRISD FLQQTTSKIY SDKEFAGSYN SSDDEVYSRP SSLVSSSPQR 60
 STSSDADGER QFNLNGNSVP TATQTLMTMY RRRSPEEFNP KLAKEFLKEQ AWKIHFAEYG 120
 QGICMYRTEK TRELVLKGIP ESMRGELWLL LSGAINEKAT HPGYYEDLVE KSMGKYNLAT 180
 EEIERDLHRS LPEHPAFQNE MGIAALRRVL TAYAFRNPNI GYCQAMNIVT SVLLLYAKEE 240
 EAFWLLVALC ERMFLPDYNT RVVGALVDQG VFEELARDYV POLYDCMQDL GVISTISLSW 300
 FLTLFLSVMP FESAVVVVDC FFYEGIKVIF QLALAVLDAN VDKLLNCKDD GEAMTVLGRY 360
 LDSVTNKDST LPPIPHLHSL LSDDVEPYPE VDFRLIRTS YEKFGTIRAD LIEQMRFKQR 420
 LKVIQTLED TCRNVVRTIV TETSFTIDEL EELYALFKAE HLTSCYWGGG SNALDRHDP 480
 LPYLEQYRID FEQFKGMFAL LPPWACGTHS DVLASRLFQL LDENGDSLIN FREFVSGLSA 540
 ACHGDLTEKL KLLYKMHVLP EPSSDQDEPD SAFEATQYFF EDITPECTHV VGLDSRSKQG 600
 ADDGFVTVSL KPDKGKRANS QENRNYLRLW TPENKSKSKN AKDLPKLNQG QFIELCKTMY 660
 NMFSEDPNEQ ELYHATAAVT SLLLEIGEVG KLFVAQPAKE GSGSGSGPSC HQGIPGVLPF 720
 KKGPGQPYVV ESVEPLPASL APDSEHSLG QMEDIKLED SSPRDNGACS SMLISDDDTK 780
 DDSSMSYSV LSAGSHEEDK LHCEEIGEDT VLVRSGQGT ALPRSTSLDR DWAITFEQFL 840
 ASLLTEPALV KYFDKPVCM ARITSANIR MMGKPLTSAS DYEISAMSG

BCU7 DNA sequence (SEQ ID NO:12)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Nucleic Acid
 Accession #: n/a; Coding sequence: 1-573 (stop codon underlined)

TATTTTATTT TCCAGGCTAA AGCAAATGAA AGTTTGCTGG TATCAACACA GCCTGCCATA 60
 TTTTTCACAG CATGCAACAA TGGTGCTAGG ATAGCTATTT CTTACTGTAA TTGCCAGAGG 120
 CAGAAATGGT CTGGGTATAA GCTATTTTCAT AAAAGCAGCT TTAAATTGTC AGTATTAAGG 180
 TTTTCATGTG GAAAGGTGTC ATTCAAAAAA AAAGTAATTG GCATACATAT TCCACATCAT 240
 CGATCCTCTC TGTGGTGTTA ATTTTATTTAT ATGACCAGTA GAAAAATTTT AATATTCTCA 300
 CAATATAGGT TTTGGGGGCTT CCATATCATC AAAAGACTGA AAAATTATAA TTTTAGAATT 360
 AAACGTATGG ATTTTCATTAT AGAATTATCT GTGAGTTGTG TAGACACAGT CTTAATGTTT 420
 CTGGTTATGA CAGATAAGTT TGCTCAAAAA ATGTGGATGA AGCCATTATT GTTATTATTG 480
 TTATTGCTTC TGTTCAGTTG TCTAAGTATC ATCCCTTCTG TGGCCCATCA CGCAGCAGAG 540
 TTGCCCTACA AATTTTCATT GGCAGCGCCA TAACATTCAT TTAATAAGTT TATGAAAACA 600
 TTCATTTGAA AGTTCCCATG AGCTTTAGCA CAGAGTTGAC CAAACACTGG CGTAAGTTCA 660
 ATTTACACAG AATATTGAA TTGAAACAAT AGAAATTTT CTCATAATAT ATACCTATGT 720
 GAAACCAACT TATCTGCATA ATTAAATCTA ATACATATTT AAGCCAGTTT AAGTGCTTTG 780
 TGTGTATGCC ATGCTTATCA AATACATGCA CAAGCTAAAC ATAATTGAA TGGGTCTATG 840
 AAGGAAAAAT AATGCTTAGA CTTTGGTGTA GGTTCCTTCT GTGTAGCCAT ATACCCAGGC 900
 TCTGCAGTAT CGAAGGATGC AAATGTTGAC ATAGATGGAA GCTCTTACCT ACCAAGTGT 960
 TTAGGAAGGA TAAAGTTACA TTTGTCTTAA TTTCTAACAT TATCTTGCT TTTATGTTT 1020
 ATAAAAATTT GTCATTATTT ATGCTGGTGA AACGTATAAT CACATCCAAT TATTTGAACA 1080
 CATGCAAAAT AATTTTAA ATTATGTTAT TGTTTAAATT TGACTTATGG GAGATCAGTC 1140
 AAAAAGTTAG AAGGTTTAA ACITCACTGA TTAATGGTGC TGAAAACACG TTACAATTAC 1200
 CACATATCCT TGCTATAAGT TTTGAAGTTT CTTAGCAATT AAAGTTTTTT TATTCAGTGT 1260
 GAACTGTCAG TATCTATTCT GGTGCTAAAT GTATGGTGCT AAATGAATTG TTAGTGTGA 1320
 TGGCTTTAGT AATGCTCCTT TTATTCATTG CTAAATTTAG TGTTATCCAT TTGATTCCTG 1380
 ATTACAGAAAT ATCAATAAAA TCCTATGTTA AATTAATCTT TACCAAAAAC AGGCAAGTTA 1440
 ACTCTGTTGT TTTAATTCAA CAGTCCAACA TTATTAGGT GTTACAGAGT GTAAATATAT 1500
 TTCTTTGGGA GTTATTTTCT TTTTAAATC TTTTATAGC TTGGCAATGT CCAAAGTCAA 1560
 ATATCACCTA AACTGGTTAG ATTACTTCTA CAGCTAATAA TATTGCAGGC ACTGGCGCCC 1620
 TCTGGTGGTT ATGAAGACAA ATTCTTAATG GCTACTTGAC CTACAGCAAA AGCCATTCT 1680
 GTACCATAAA AATTTGTTGT GCAATATTAG AATTATCATA TGTTTCCTAC ATCTGACAGC 1740
 ACCTAAAAAT TTTGATAATA TTAACATGTA TCTAAGAGGA AAAAAAGATT AATATATTCT 1800
 GGCACCCACT TTCCTAGTAA TGTTTTCCAT GATTTTCCAG TTCTGAGGCA CTTATTAAAG 1860
 TGCTTTTTTT TTCTGAATT AATTAGGTAT TGGTAAATA TATTTTAA TTTAGTTAGC 1920
 TTTATAAACA CTTTGAAT TACAATTAA TAACAGAGGT ATAATTGTCT CACTTTCAGA 1980
 AGTGATCATT TATTTTATT TAGCACAGGT CATAAGAAAA ATATATAGAA AAATAATCAA 2040

```

TTTCATATAT AAAAGGATTA TTTCTCCACC TTTAATTATT GGCCTATCAT TTGTTAGTGT 2100
TATTTGGTCA TATTATTGAA CTAATGTATT ATTCCATTCA AAGTCTTTCT AGATTTAAAA 2160
ATGTATGCAA AAGCTTAGGA TTATATCATG TGTAACATATT ATAGATAACA TCCTAAACCT 2220
TCAGTTTAGA TATATAATTG ACTGGGTGTA ATCTCTTTTG TAATCTGTTT TGACAGATTT 2280
CTTAAATTAT GTTAGCATAA TCAAGGAAGA TTTACCTTGA AGCACTTTCC AAATTGATAC 2340
TTTCAAACCT ATTTTAAAGC AGTAGAACCT TTTCTATGAA CTAAATCACA TGCAAAATC 2400
CAACCTGTAG TATACATAAA ATGGACTTAC TTATTCTCTC CACCTTCTCC AGTGCCTAGG 2460
AATATTCTTC TCTGAGCCCT AGGATTGATT CTATCACACA GAGCAACATT AATCTAAATG 2520
GTTTAGCTCC CTCTTTTTC TCTAAAAACA ATCAGCTAAT AAAAAAATAA TTTGAGGGCC 2580
TAAATTATTT CAATGGTTGT TTGAAATATT CAGTTCAGTT TGTACCTGTT AGCAGTCTTT 2640
CAGTTTGGGG GAGAATTAAA TACTGTGCTA AGCTGGTGCT TGGATACATA TTACAGCATC 2700
TTGTGTTTTA TTTGACAAAC AGAATTTTGG TGCCATAATA TTTTGAGAAT TAGAGAAGAT 2760
TGTGATGCAT ATATATAAAC ACTATTTTAA AAAAATATCT AAATATGTCT CACATATTTA 2820
TATAATCCTC AAATATACTG TACCATTTTA GATATTTTTT AAACAGATTA ATTTGGAGAA 2880
GTTTTATTCA TTACCTAATT CTGTGGCAA AATGGTGCCT CTGATGTTGT GATATAGTAT 2940
TGTCAGTGTG TACATATATA AAACCTGTGT AAACCTCTGT CCTATGAAC CATAACAAAT 3000
GTAGCTTTTT AAAGTCCATT GTATTGTTTT TTCTTTCAAT AAAAGAGTAT AATTAATTGG 3060
TTGTTTTTGA

```

BCU7 Protein sequence (SEQ ID NO:13)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 125-141, 154-170; PFAM domains: none; Summary: A type III membrane protein, highly overexpressed in breast cancer and prostate cancer; unknown function.

```

YFIFQAKANE SLLVSTQPAI FFTACNNGAR IAISYCNCQR QKWSGYKLFH KSSFKLSVLR 60
FSCGKVSFKK KVGIGIHPHH RSSLWCXFFY MTSRKILIFS QYRFWGFHII KRLKNYNFRI 120
KLMDFIIELS VSCVDTVLMF LVMTDKFAQK MWMKPLLLLL LLLLFSCLSI IPSVAHHAAE 180
LPYKFHLAAP

```

BFA1 DNA sequence (SEQ ID NO:14)

Gene name: calsynenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Nucleic Acid Accession #: NM_022131; Coding sequence: 11-2878 (start and stop codons underlined)

```

TGCTGCGAGG ATGCTGCCTG GCGGCTGTG CTGGGTGCCG CTCCTGCTGG CGCTGGGCGT 60
GGGGAGCGGC AGCGGCGGTG GCGGGGACAG CCGGCAGCGC CGCCTCCTCG CGGCTAAAGT 120
CAATAAGCAC AAGCCATGGA TCGAGACTTC ATATCATGGA GTCATAACTG AGAACAATGA 180
CACAGTCATT TTGGACCCAC CACTGGTAGC CCTGGATAAA GATGCACCGG TTCCTTTTGC 240
AGGGGAAATC TGTGCGTTCA AGATCCATGG CCAGGAGCTG CCCTTTGAGG CTGTGGTGCT 300
CAACAAGACA TCAGGAGAGG GCCGGCTCCG TGCCAAGAGC CCCATTGACT GTGAGTTGCA 360
GAAGGAGTAC ACATTCAATCA TCCAGGCCTA TGACTGTGGT GCTGGGCCCC ACGAGACAGC 420
CTGGAAAAAG TCACACAAGG CCGTGGTCCA TATACAGGTG AAGGATGTCA ACGAGTTTGC 480
TCCCACCTTC AAAGAGCCAG CCTACAAGGC TGTGTGACG GAGGGCAAGA TCTATGACAG 540
CATTCTGCAG GTGGAGGCCA TTGACGAGGA CTGCTCCCCA CAGTACAGCC AGATCTGCAA 600
CTATGAAATC GTCACCACAG ATGTGCCTTT TGCCATCGAC AGAAATGGCA ACATCAGGAA 660
CACTGAGAAG CTGAGCTATG ACAAACAACA CCAGTATGAG ATCCTGGTGA CCGCCTACGA 720
CTGTGGACAG AAGCCCGCTG CTCAGGACAC CCTGGTGACG GTGGATGTGA AGCCAGTTTG 780
CAAGCCTGGC TGGCAAGACT GGACCAAGAG GATTGAGTAC CAGCCTGGCT CCGGGAGCAT 840
GCCCCTGTTT CCCAGCATCC ACCTGGAGAC GTGCGATGGA GCCGTGTCTT CCCTCCAGAT 900
CGTCACAGAG CTGCAGACTA ATTACATTGG GAAGGGTTGT GACCGGGAGA CCTACTCTGA 960
GAAATCCCTT CAGAAGTTAT GTGGAGCCTC CTCTGGCATC ATTGACCTCT TGCCATCCCC 1020
TAGCGCTGCC ACCAACTGGA CTGCAGGACT GCTGGTGGAC AGCAGTGAGA TGATCTTCAA 1080
GTTTTCAGGC AGGCAGGGTG CCAAAATCCC CGATGGGATT GTGCCCCAAGA ACCTGACCGA 1140
TCAGTTTACC ATCACCATGT GGATGAAACA CGGCCCCAGC CCTGGTGTGA GAGCCGAGAA 1200
GGAAACCATC CTCTGCAACT CAGACAAAAC CGAAATGAAC CGGCATCACT ATGCCCTGTA 1260
TGTGCACAAC TGCCGCTCGT TCTTTCTCTT GCGGAAGGAC TTCGACCAGG CTGACACCTT 1320
TCGCCCCGCG GAGTTTCCACT GGAAGCTGGA TCAGATTTGT GACAAAGAGT GGCACTACTA 1380
TGTCATCAAT GTGGAGTTTC CTGTGGTAAC CTTATACATG GATGGAGCAA CATATGAACC 1440
ATACCTGGTG ACCAACGACT GGCCCATTC A TCCATCTCAC ATAGCCATGC AACTCACAGT 1500
CGGCGCTTGT TGGCAAGGAG GAGAAGTCAC CAAACCACAG TTTGCTCAGT TCTTTCATGG 1560
AAGCCTGGCC AGTCTCACC TCCGCCCTGG CAAAATGGAA AGCCAGAAGG TGATCTCCTG 1620
CCTGCAGGCC TGCAAGGAAG GGCTGGACAT TAATTCCTTG GAAAGCCTTG GCCAAGGAAT 1680
AAAGTATCAC TTCAACCCCT CGCAGTCCAT CCTGGTGATG GAAGGTGACG ACATTGGGAA 1740
CATTAACCGT GCTCTCCAGA AAGTCTCCTA CATCAACTCC AGGCAGTTCC CAACGGCGGG 1800
TGTGCGGCGC CTCAAAGTAT CCTCCAAAGT CCAGTGCTTT GGGGAAGACG TATGCATCAG 1860
TATCCCTGAG GTAGATGCCT ATGTGATGGT CCTCCAGGCC ATCGAGCCCC GGATCACCTT 1920
CCGGGGCACA GACCACTTCT GGAGACCTGC TGCCCAGTTT GAAAGTGCCA GGGGAGTGAC 1980
CCTCTTCCCT GATATCAAGA TTGTGAGCAC CTTGCCCCAA ACCGAAGCCC CCGGGGACGT 2040
GAAAACCACA GACCCCAAT CAGAAGTCTT AGAGGAAATG CTTCACTAAT TAGATTTCTG 2100
TGACATTTTG GTGATCGGAG GGGACTTGGA CCAAGGCAG GAGTGCTTGG AGCTCAACCA 2160
CAGTGAGCTC CACCAACGAC ACCTGGATGC CACTAATTCT ACTGCAGGCT ACTCCATCTA 2220
CGGTGTGGGC TCCATGAGCC GCTATGAGCA GGTGCTACAT CACATCCGCT ACCGCAACTG 2280
GCGTCCGGCT TCCCTTGAGG CCCGGCGTTT CCGGATTAAG TGCTCAGAAC TCAATGGGCG 2340
CTACAGTACG AATGAGTTCA ACTTGGAGGT CAGCATCCTT CATGAAGACC AAGTCTCAGA 2400
TAAGGAGACT GTCAATCTTC TGATTGTGCA GCCTCCCTTC CTCCAGTCTG TCCATCATCC 2460
TGAGTCCCGG AGTAGCATCC AGCACAGTTC AGTGGTCCCA AGCATTGCCA CAGTGGTCAT 2520

```


CATCATCTCC	GTGTGCATGC	TTGTGTTTGT	CGTGGCCATG	GGTGTGTACC	GGGTCCGGAT	2580
CGCCACCAG	CACCTTCATCC	AGGAGACTGA	GGCTGCCAAG	GAATCTGAGA	TGGACTGGGA	2640
CGATTCTGCG	CTGACTATCA	CAGTCAACCC	CATGGAGAAA	CATGAAGGAC	CAGGGCATGG	2700
GGAAGATGAG	ACTGAGGGAG	AAGAGGAGGA	AGAAGCCGAG	GAAGAAATGA	GCTCCAGCAG	2760
TGGCTCTGAC	GACAGCGAAG	AGGAGGAGGA	GGAGGAAGGG	ATGGGCAGAG	GCAGACATGG	2820
GCAGAATGGA	CCCAGGCAAG	CCCAGCTGGA	GTGGGATGAC	TCCACCTCC	CCTACTAGTG	2880
CCCAGGGGTC	TGCTGCCTGG	CCCACATGTC	CCTTTTGTA	ACCCTGACCC	AGTGTATGCC	2940
CATGTCTATC	ATACCTCACC	TCTGATGTCT	GTGACATGTC	TGGGAAGGCC	TTCTCCAGCT	3000
TCCTGGAGCC	CACCCTTAA	GCCTTGGCA	CTCCCTGTGT	TTCATCCATG	GGGAAGTTCC	3060
AAGAAGCCCA	GCATGGCCAT	CAGTGAGGAC	TTCAGGGTAG	ACTTTGTCTT	GTAGCCTCCA	3120
CTTCTGCCCT	AAGTTCCTCA	GCATCCTGAC	TACCTGTCTG	CAGAGTTTGC	CTTTGTTTTT	3180
TCCTGCAGGG	AAGAAGGCC	ACCTTTGTGT	CACTCACCTC	CCCAGGCTCA	GAGTCCCAA	3240
GGCCCTGGGG	TTCCAACCTA	CTGTGCGTCT	CCTCCACACA	GACCAGTAGG	TTCTCCTATG	3300
CTGACTCCAG	GTTGCTTCAT	ACAAGGAGGG	TGGTTGAACT	TCACACACGT	AAGGTCTTAG	3360
TGCTTAACAG	TTTAAAGGAA	AGTCCTTGTT	GAGGCAGAAC	TAAGTTTACA	GGGAAAGGTA	3420
CACACATTCT	CTCTCTCTCT	CTCTCTCTGT	CTATCTAGTT	CCCCAGCTTG	GAGAGCCTTT	3480
CCCCTTGCTT	CTTTCTGAGG	CCATATAAGC	TTATAAGAAA	AGTCCCAAAC	CAAGAATAGG	3540
TCCTTGGCCA	CAAGCAGGGT	CTGATCCCCC	ATCAGAGCTA	TCTGAGCCTG	CCTGTCTGGG	3600
CACCTGCTGC	AACCATGCAG	CTACCCTGCC	AGGGGCACTC	AGCAAACAGA	ACCACAGGGC	3660
CCAGGAGGCA	TTCCACACAG	GCATGCCCCC	AGGACAACAC	AACAAGGACA	GTCACAACAA	3720
GGACAACAAG	GACACAACAC	AACACAACAC	AAGGACAGTC	ACAACAAGCC	TAGAGCCAGA	3780
AAGCAGATGG	AAATGCTAAT	GAGGTCAAAC	GTAGGCTTCA	TGGTGGGTGG	AGTGGGGTGG	3840
GCTGGGCTCC	CCCAGGACAG	AGGGGACCTT	GAGGTTGGCA	AGGCTCTCAC	CACCTAGCCT	3900
TATGGTCCCT	TATCTCCTAT	CTTCCCTCTT	GAGAAAATAC	ACGCTTTCTG	CATGTATTAG	3960
AAACGCACGA	GCTCCACCAA	GTCTACAATG	AAAGTTTGAA	ATTAACTG	AAGGAATTAG	4020
AAGCATATTT	GCAATCATTG	CAGCTTCTTC	TTCTTCTG	TCATAAAAGG	AGGAACACTT	4080
TAGATAGAGG	GCAAAATATAT	CTGAAAACCT	AATTTCTTTC	TTTTTTTGAT	AAGGAAATCT	4140
TTTCCATCTC	CATCCTAACA	TGCACAACCT	GTGAAGAGAA	TTGTTTCTAT	AGTAACTGGT	4200
CTGTGATCTT	TGTGGGCCAA	GAGAATAGCA	GGCAAGAATT	AGGGCCTTGA	CAGAATTTCC	4260
ACGAAGCTCT	GAGAACATGT	TGTTTTCGAA	TGCTGATTC	CTCTTTGTCA	TCAATGTGTA	4320
TGCTCTGTCC	CCATCCTTCA	CTCCTCCTCA	AGCTCACACC	AATTGGTTTG	GCACAGGCAC	4380
AGAGCTGGTC	CCTAGTTAAG	TGGCATTAT	GTAAAAAAA	A		

BFA1 Protein sequence (SEQ ID NO:15)

Gene name: calsynenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Protein Accession #: NP_071414; Predicted Signal sequence: 1-20; Predicted TM domains: 832-848; PFAM domains: cadherin domains: 48-151, 165-254; Summary: A type I membrane protein; a member of the calsynenin family; is related to the FAT tumor suppressor; is likely an adhesion molecule important in mammalian developmental processes and cell communication.

MLPGRLCWVP	LLALGVGSG	SGGGDSRQR	RLLAQVKNKH	KPWIETSYHG	VITENNDTVI	60
LDPPLVALDK	DAPVPFAGEI	CAFKIHQEL	PFEAVVLNKT	SGEGRRLRAKS	PIDCELQKEY	120
TFIIQAYDCG	AGPHETAWKK	SHKAVVHIQV	KDVNEFAPTF	KEPAYKAVT	EGKIYDSILQ	180
VEAIDEDCSP	QYSQICNYEI	VTTDVPFAID	RNGNIRNTEK	LSYDKQHQYE	ILVTAYDCGQ	240
KPAAQDTLVQ	VDVKPVCKPG	WQDWTKRIEY	QPGSGSMPLE	PSIHLETCDG	AVSSLQIVTE	300
LQTNYYIGKC	DRETYSEKSL	QKLCGASSGI	IDLLPSPSAA	TNWTAGLLVD	SSEMIFKFDG	360
RQAKIPDGI	VPKNLTDQFT	ITMWMKHGPS	PGVRAEKETI	LCNSDKTEMN	RHHYALYVHN	420
CRLVFLLRKD	FDQADTFRPA	EFHWKLDQIC	DKEWHYYVIN	VEFPVVTLYM	DGATYEPYLV	480
TNDWPIHPSH	IAMQLTVGAC	WQGEVTKPQ	FAQFFHGSLS	SLTIRPGKME	SQKVISCLQA	540
CKEGLDINSL	ESLQGQIKYH	FNPSQSILVM	EGDDIGNINR	ALQKVSYYNS	RQFPPTAGVRR	600
LKVSQKQCF	GEDVCISIEP	VDAYVMVLQA	IEPRITLRGT	DHFWRPAAQF	ESARGVTLFP	660
DIKIVSTFAK	TEAPGDVKT	DPKSEVLEEM	LHNLDFCIL	VIGGDLDPQ	ECLELNHSEL	720
HQRHLDATNS	TAGYSIYGVG	SMSRYEQVLH	HIRYRNWRPA	SLEARRFRIK	CSELNRYTS	780
NEFNLEVSIL	HEDQVSDKEH	VNHLIVQPPF	LQSVHHPESR	SSIQHSSVVP	SIATVVIIIS	840
VCMLVFFVAM	GVYRVRIAHQ	HFIQETEA	ESEMWDWDSA	LTITVNPMEK	HEGPGHGEDE	900
TEGEEEEAE	EEMSSSSGSD	DSEEEEEEEG	MGRGRHGQNG	ARQAQLEWDD	STLPY	

BFG7 DNA sequence (SEQ ID NO:16)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Nucleic Acid Accession #: n/a; Coding sequence: <1-906 (stop codon underlined)

CGGGTCGACC	CACGCGTCCG	GGGAGAAAGG	ATGGCCGGCC	TGGCGGCGCG	GTTGGTCCTG	60
CTAGCTGGGG	CAGCGCGCCT	GGCGAGCGGC	TCCCAGGGCG	ACCGTGAGCC	GGTGTACCGC	120
GACTGCGTAC	TGCAGTGCGA	AGAGCAGAAC	TGCTCTGGGG	GCGCTCTGAA	TCACTTCCGC	180
TCCCGCCAGC	CAATCTACAT	GAGTCTAGCA	GGCTGGACCT	GTCGGGACGA	CTGTAAGTAT	240
GAGTGTATGT	GGGTACCCGT	TGGGCTCTAC	CTCCAGGAAG	GTCACAAAGT	GCCTCAGTTC	300
CATGGCAAGT	GGCCCTTCTC	CCGGTTCCTG	TTCTTTCAAG	AGCCGGGCATC	GGCCGTGGCC	360
TCTTTTCTCA	ATGGCTGGG	CAGCCTGGTG	ATGCTCTGCC	GCTACCGCAC	CTTCGTGCCA	420
GCCTCCTCCC	CCATGTACCA	CACCTGTGTG	GCCTTCGCCT	GGGTGTCCCT	CAATGCATGG	480
TTCTGGTCCA	CAGTYTTCCA	CACCAGGGAC	ACTGACCTCA	CAGAGAAAAT	GGACTACTTC	540
TGTGCCTCCA	CTGTATCCT	ACACTCAATC	TACCTGTGCT	GCGTCAGCCT	CATCCGCTTC	600
GACTATGGCT	ACAACCTGGT	GGCCAACGTG	GCTATTGGCC	TGGTCAACGT	GGTGTGGTGG	660
CTGGCCTGGT	GCCTGTGGAA	CCAGCGGCGG	CTGCCTCACG	TGCGCAAGTG	CGTGGTGGTG	720
GTCTTGCTGC	TGCAGGGGCT	GTCCCTGCTC	GAGCTGCTTG	ACTTCCCACC	GCTCTTCTGG	780
GTCCTGGATG	CCCATGCCAT	CTGGCACATC	AGCACCATCC	CTGTCCACGT	CCTCTTTTTC	840
AGCTTTCTGG	AAGATGACAG	CCTGTACCTG	CTGAAGGAAT	CAGAGGACAA	GTTCAAGCTG	900

GACTGAAGAC CTTGGAGCGA GTCTGCCCCA GTGGGGATCC TGCCCCCGCC CTGCTGGCCT 960
 CCCTTCTCCC CTC AACCTT GAGATGATTT TCTCTTTTCA ACTTCTTGAA CTTGGACATG 1020
 AAGGATGTGG GCCCAGAATC ATGTGGCCAG CCCACCCCTC GTTGGCCCTC ACCAGCCTTG 1080
 GAGTCTGTTC TAGGGAAGGC CTCCAGCAT CTGGGACTCG AGAGTGGGCA GCCCCTCTAC 1140
 CTCCTGGAGC TGAAGTGGG TGGAAGTGGG TGTGCTCTTA GCTCTACCGG GAGGACAGCT 1200
 GCCTGTTTCC TCCCCATCAG CCTCTCCCC ACATCCCCAG CTGCCTGGCT GGGTCCCTGAA 1260
 GCCCTCTGTC TACCTGGGAG ACCAGGGACC ACAGGCCTTA GGGATACAGG GGGTCCCTT 1320
 CTGTTACCAC CCCCCACCCT CCTCCAGGAC ACCACTAGGT GGTGCTGGAT GCTTGTCTTT 1380
 TGGCCAGCCA AGGTTACAGG CGATTCTCCC CATGGGATCT TGAGGGACCA AGCTGCTGGG 1440
 ATTGGGAAGG AGTTTACCC TGACCRITGC CCTAGCCAGG TCCCCAGGAG GCCTCACCAT 1500
 ACTCCCTTTC AGGGCCAGGG CTCCAGCAAG CCCAGGGCAA GGATCCTGTG CTGCTGTCTG 1560
 GTTGAGAGCC TGCCACCGTG TGTGCTGGAG GTGGGCCAGG CTGAGTGCAT AGGTGACAGG 1620
 GCCGTGAGCA TGGGCTGGG TGTGTGTGAG CTCAGGCACT AGGTGCGCAG TGTGGAGACG 1680
 GGTGTGTGCG GGGGAAGAGT GTGGCTCAA AGTGTGTGTG GTGCAGGGGG TKGGTGTGTT 1740
 AAGCGTGGGT TAGGGGAACG TGTGTGCGCG TGCTGGTGGG CATGTGAGAT GAGTGAAGTGC 1800
 CGGTGAATGT GTCCACAGTT GAGAGGTTGG AGCAGGATGA GGAATCCTG TCACCATCAA 1860
 TAATCACTTG TGGAGCGCCA CTTGGCCCAA GACGCCACCT GGGCGGACAG CAGGAGCTCT 1920
 CCATGGCCAG GCTGCTGTG TGCATGTTCC CTGTCTGGTG CCCCTTTGCC CGCCTCCTGC 1980
 AAACCTCACA GGGTCCCCAC ACAACAGTGC CCTCCAGAAG CAGCCCTCG GAGGCAGAGG 2040
 AAGGAAAATG GGGATGGCTG GGGCTCTCTC CATCTCTCTT TTCTCTTGC CTTCGCATGG 2100
 CTGGCCCTCC CCTCCAAAAC CTCCATCCCC CTGCTGCCAG CCCCTTTGCC ATAGCCTGAT 2160
 TTTGGGGAGG AGGAAGGGGG GATTTGAGGG AGAAGGGGAG AAAGCTTATG GCTGGGTCTG 2220
 GTTTCTTCCC TTCCACAGAG GTCTTACTGT TCCAGGGTGG CCCCAGGCA GCAGGGGCC 2280
 AACTATGCC TGCGCCCTGG TAAAGGTGAC CCCTGCCATT TACCAGCAGC CTTGGCATGT 2340
 TCCTGCCCCA CAGGAATAGA ATGGAGGGAG CTCCAGAAAC TTTCATCCC AAAGGCAGTC 2400
 TCCGTGGTTG AAGCAGACTG GATTTTGTCT CTGCCCTGA CCCCTTGTC CTCTTTGAGG 2460
 GAGGGGAGCT ATGCTAGGAC TCCAACCTCA GGGACTCGGG TGGCCTGCGC TAGCTTCTTT 2520
 TGATACTGAA AACTTTTAAG GTGGGAGGGT GGCAAGGGAT GTGCTTAATA AATCAATTCC 2580
 AAGCCTCAAA AAAAAAAAAA AAAAAAAAAA AAAAAA

BFG7 Protein sequence (SEQ ID NO:17)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Protein Accession
 #: n/a; Signal sequence: none; Predicted TM domains: 117-133, 179-195, 211-227, 235-251,
 266-282, 296-312; PFAM domains: none; Summary: A type III membrane protein of unknown
 function; is adjacent to HER2 on the genome, and its overexpression in breast cancer is highly
 correlated with HER2 expression; may be used to predict HER2 overexpression and amplification.

RVDPRVRGER MAGLAARLVL LAGAAALASG SQGDREPVR DCVLQCEEQN CSGGALNHFR 60
 SRQPIYMSLA GWTCRDDCKY ECMWVTVGLY LQEGHKVPQF HGKWPFSRFL FFQEPASAVA 120
 SFLNGLASLV MLCRYRTFVP ASSPMYHTCV AFAWVSLNAW FWSTVFHTRD TDLTEKMDYF 180
 CASTVILHSI YLCCVRTVGL QHPAVVSAFR ALLLLMLTVH VSYLSLIRFD YGYNLVANVA 240
 IGLVNVVWVL AWCWLNQRRRL PHVRKCVVWV LLLQGLSLE LLDFFPLFWV LDAHAIWHIS 300
 TIPVHVLFFS FLEDDSLYLL KESEDKFKLD

BCN4 DNA sequence (SEQ ID NO:18)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Nucleic Acid
 Accession #: n/a; Coding sequence: 143-874 (start and stop codons underlined)

GGGAGGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCCTCGGAG 60
 CGCGGCGGAG CCAGACGCTG ACCACGTTCC TCTCTCGGT CTCCTCCGCC TCCAGCTCCG 120
 CGCTGCCCGG CAGCCGCGGACC CCATGCGACC CCAGGGCCCC GCGGCTCCC CGCAGCGGCT 180
 CCGCGGCCCTC CTGCTGCTCC TGCTGCTGCA GCTGCCCGCG CCGTCGAGCG CCTCTGAGAT 240
 CCCCAAGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG 300
 AATGTGCTTA CAAGGGCCAG CAGGAGTGCC TGGTCGAGAC GGGAGCCCTG GGGCCAATGG 360
 CATTCGGGT ACACCTGGGA TCCCAGGTCG GGATGGATTG AAAGGAGAAA AGGGGAATG 420
 TCTGAGGGAA AGCTTTGAGG AGTCCTGGAC ACCCAACTAC AAGCAGTGTT CATGGAGTTC 480
 ATTGAATTAT GGCATAGATC TTGGGAAAAT TGCGGAGTGT ACATTACAA AGATGCGTTC 540
 AAATAGTGCT CTAAGAGTTT TGTTCACTGG CTCACCTCGG CTAAATGCA GAAATGCATG 600
 CTGTCAGCGT TGGTATTTCA CATTCAATGG AGCTGAATGT TCAGGACCTC TTCCCATGTA 660
 AGCTATAATT TATTTGGACC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTCATCG 720
 CACTTCTTCT GTGGAAGGAC TTTGTGAAGG AATTGGTGCT GGATTAGTGG ATGTTGCTAT 780
 CTGGGTGGC ACTTGTTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT 840
 TTCTCGCATC ATTATTGAAG AACTACCAAA ATAAATGCTT TAATTTTCAT TTGCTACCTC 900
 TTTTATTATT ATGCTTGGG ATGGTTCAT TAAATGACAT TTTAAATAAG TTTATGTATA 960
 CATCTGAATG AAAAGCAAAG CTAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT 1020
 TTTAAATCTA GCATTATTCA TTTTGCTTCA ATCAAAAGTG GTTTCAATAT TTTTCTTAGT 1080
 TGGTTAGAAAT ACTTCTTCA TAGTCACATT CTCTCAACCT ATAATTGGA ATATTGTGTG 1140
 GGTCTTTTGT TTTTCTCTT AGTATGCAT TTTTAAAAA ATATAAAGC TACCAATCTT 1200
 TGTACAATTT GTAAATGTTA AGAATTTTTT TTATATCTGT TAAATAAAAA TTATTTCCAA 1260
 CAACCTTAAA AAAAAAAAAA AAAAA

BCN4 Protein sequence (SEQ ID NO:19)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Protein Accession
 #: n/a; Predicted Signal sequence: 1-30; TM domains: none; PFAM domains: none; Summary: a
 secreted protein; has a mouse orthologue (see sequence below).

MRPQGPAASP	QRLRGLLLLL	LLQLPAPSSA	SEIPKGKQKA	QLRQREVVDL	YNGMCLQGPA	60
GVPGRDGSPG	ANGIPGTPGI	PGRDGFKEG	GECLRESFEE	SWTPNYKQCS	WSSLNYGIDL	120
GKIAECTFTK	MRSNSALRVL	FSGSLRLKCR	NACCQRWYFT	FNGAECGGL	PIEAIYLDQ	180
GSPEMNSTIN	IHRTSSVEGL	CEGIGAGLVD	VAIWVGTCSD	YPKGDASTGW	NSVSRIIEE	240

LPK

Mouse BCN4 Protein sequence (SEQ ID NO:20)
Gene name: ESTs; Unigene number: Mm.41556

XXXXAAPPQL	LLGLFLVLLL	LLQLSAPSSA	SEIPKVKQKA	LIRQREVVDL	YNGMCLQGPA	60
GVPGRDGSPG	ANGIPGTPGI	PCQDGFKEG	GECLRESFEE	SWTPNYKQCS	WSSLNYGIDL	120
GKIAECTFTK	MRSNSALRVL	FSGSLRLKCR	NACCQRWYFT	FNGAECGGL	PIEAIXXXXX	180
XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXSD	YPKGDAYTGW	DSVSRIIEE	240

LPK

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Specification:

Paragraph (Table 1) beginning at line 1 of page 94 has been amended as follows (see attached pages 94-103):

Table 1

BCA4 DNA sequence (SEQ ID NO:1)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset
Accession #: D13666; Nucleic Acid Accession #: NM_006475; Coding sequence: 12-2522 (start
and stop codons underlined)

```

AGAGACTCAA GATGATTCCC TTTTACCCA TGTCTCTCT ACTATTGCTG CTTATTGTTA 60
ACCCATATAA CGCCAACAAT CATTATGACA AGATCTTGGC TCATAGTCGT ATCAGGGGTC 120
GGGACCAAGG CCCAAATGTC TGTGCCCTTC AACAGATTTT GGGCACCAAA AAGAAATACT 180
TCAGCACTTG TAAGAACTGG TATAAAAAGT CCATCTGTGG ACAGAAAACG ACTGTTTTAT 240
ATGAATGTTG CCCTGGTTAT ATGAGAATGG AAGGAATGAA AGGCTGCCCA GCAGTTTTGC 300
CCATTGACCA TGTTTATGGC ACTCTGGGCA TCGTGGGAGC CACCACAACG CAGCGCTATT 360
CTGACGCCTC AAAAATGAGG GAGGAGATCG AGGGAAAGGG ATCCTTCACT TACTTTGCAC 420
CGAGTAATGA GGCTTGGGAC AACTTGGATT CTGATATCCG TAGAGGTTTG GAGAGCAACG 480
TGAATGTTGA ATTACTGAAT GCTTTACATA GTCACATGAT TAATAAGAGA ATGTTGACCA 540
AGGACTTAAA AAATGGCATG ATTATTCCTT CAATGTATAA CAATTTGGGG CTTTTCATTA 600
ACCATTTATC TAATGGGGTT GTCACGTGTA ATTGTGCTCG AATCATCCAT GGGAAACCAGA 660
TTGCAACAAA TGGTGTGTGC CATGTGATTG ACCGTGTGCT TACACAAATT GGTACCTCAA 720
TTCAAGACTT CATTGAAGCA GAAGATGACC TTTTCATCTT TAGAGCAGCT GCCATCACA 780
CGGACATATT GGAGGCCCTT GGAAGAGACG GTCACCTCAC ACTCTTTGCT CCCACCAATG 840
AGGCTTTTGA GAAACTTCCA CGAGGTGTCC TAGAAAGGTT CATGGGAGAC AAAGTGGGCTT 900
CCGAAGCTCT TATGAAGTAC CACATCTTAA ATACTCTCCA GTGTTCTGAG TCTATTATGG 960
GAGGAGCAGT CTTTGGAGCG CTGGAAGGAA ATACAATTGA GATAGGATGT GACGGTGACA 1020
GTATAACAGT AAATGGAATC AAAATGGTGA ACAAAAAGGA TATGTGACA AATAATGGTG 1080
TGATCCATTT GATTGATCAG GTCCTAATTC CTGATTCTGC CAAACAAGTT ATTGAGCTGG 1140
CTGGAACAAC GCAACCACC TTCACGGATC TTGTGGCCCA ATTAGGCTTG GCATCTGCTC 1200
TGAGGCCAGA TGGAGAATAC ACTTTGCTGG CACCTGTGAA TAATGCATT TCTGATGATA 1260
CTCTCAGCAT GGTTCAGCGC CTCCTTAAAT TAATTCTGCA GAATCACATA TTGAAAGTAA 1320
AAGTTGGCCT TAATGAGCTT TACAACGGGC AAATACTGGA AACCATCGGA GGCACACG 1380
TCAGAGTCTT CGTATATCGT ACAGCTGTCT GCATTGAAAA TTCATGCATG GAGAAAGGGA 1440
GTAAGCAAGG GAGAAACGGT GCGATTCACT TATCCGCGA GATCATCAAG CCAGCAGAGA 1500
AATCCCTCCA TGAAAAGTTA AAACAAGATA AGCGCTTTAG CACCTTCCTC AGCCTACTTG 1560
AAGCTGCAGA CTTGAAAGAG CTCCTGACAC AACCTGGAGA CTGGACATTA TTTGTGCCAA 1620
CCAATGATGC TTTTAAGGGA ATGACTAGTG AAGAAAAAGA AATTCTGATA CGGGACAAAA 1680
ATGCTCTTCA AAACATCATT CTTTATCACC TGACACCAGG AGTTTTTATT GGAAAGGAT 1740
TTGAACCTGG TGTTACTAAC ATTTTAAAGA CCACACAAGG AAGCAAAATC TTTCTGAAAG 1800
AAGTAAATGA TACACTTCTG GTGAATGAAT TGAATCAAAA AGAATCTGAC ATCATGACAA 1860
CAAATGGTGT AATTCATGTT GTAGATAAAC TCCTCTATCC AGCAGACACA CCTGTGGAA 1920
ATGATCAACT GCTGGAATA CTTAATAAAT TAATCAAATA CATCCAAAT AAGTTTGTTT 1980
GTGGTAGCAC CTTCAAAGAA ATCCCGTGA CTGTCTATAC AACTAAAATT ATAACCAAAG 2040
TTGTGGAACC AAAAATTAAA GTGATTGAA GCAGTCTTCA GCCTATTATC AAACTGAAG 2100
GACCCACACT AACAAAGTC AAAATTGAAG GTGAACCTGA ATTACAGACTG ATTAAGAAG 2160
GTGAAACAAT AACTGAAGTG ATCCATGGAG AGCCAATTAT TAAAAAATAC ACCAAATCA 2220
TTGATGGAGT GCTGTGGAA ATAACTGAAA AAGAGACACG AGAAGAACGA ATCATTACAG 2280
GTCTGAAAT AAAATACACT AGGATTTCTA CTGGAGGTGG AGAAACAGAA GAAACTCTGA 2340
AGAAATTGTT ACAAGAAGAG GTCACCAAGG TCACCAAATT CATTGAAGGT GGTGATGGTC 2400
ATTTATTTGA AGATGAAGAA ATTTAAAGAC TGCTTCAGGG AGACACACC GTGAGGAAGT 2460
TGCAAGCCAA CAAAAAAGTT CAAGTTCTTA GAAGACGATT AAGGGAAGGT CGTTCTCAGT 2520
GAAAATCCAA AAACCAGAAA AAAATGTTTA TACAACCTTA AGTCAATAAC CTGACCTTAG 2580
AAAATGTGTA GAGCAAGTT GACTTCAGGA ACTGAAACAT CAGCACAAAG AAGCAATCAT 2640
CAATAATTC TGAACACAAA TTTTATATTT TTTTCTCTGA ATGAGAAACA TGAGGGAAT 2700
TGTGGAGTTA GCCTCCTGTG GTAAAGGAAT TGAAGAAAT ATAACACCTT ACACCCTTTT 2760
TCATCTTGAC ATTAAGAGTT CTGGCTAACT TTGGAATCCA TTAGAGAAAA ATCCTTGTC 2820
CCAGATTATG TACAATTCAT ATCGAAGAGT TGTGAACTGT TATCCCATTT AAAAGACCGA 2880
GCCTTGATAT TATGTTATGG ATACATAAAA TGCACGCAAG CCATTATCTC TCCATGGGAA 2940
GCTAAGTTAT AAAAATAGGT GCTTGGTGTA CAAAACCTTT TATATCAAAA GGCTTTGCAC 3000
ATTTCTATAT GAGTGGGTTT ACTGGTAAAT TATGTTATTT TTTACAATA ATTTGTACT 3060
CTCAGAATGT TTGTCATATG CTTCTTGCAA TGCATATTTT TTAATCTCAA ACGTTTCAAT 3120
AAAACCATTT TTCAGATATA AAGAGAATTA CTTCAAATG AGTAATTCAG AAAAATCAA 3180
GATTTAAGTT AAAAAGTGGT TTGGACTTGG GAA

```

BCA4 Protein sequence (SEQ ID NO:2)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset
Accession #: D13666; Protein Accession #: NP_006466; Predicted Signal sequence: 1-21; TM
domains: none; PFAM domains: fasciclin domains: 94-232, 234-367, 496-630; Summary: a
secreted protein involved in adhesion and osteoblast development; may participate in
preferential metastasis of breast cancer to bone.

```

MIPFLPMFSL LLLLVNPIANNNHYDKILA HSRIRGRDQG PNVCALQQIL GTKKKYFSTC 60
KNWYKKSICG QKTTVLVECC PGYMRMEGMK GCPAVLPIDH VYGLTGIVGA TTTQRYSDAS 120
KLREEIEGKG SFTYFAPSNE AWDNLDSIR RGLSNVNVE LLNALHSHMI NKRLMTKDLK 180
NGMIIPSMYN NLGLFINHYP NGVVTVNCAR IIHGNQIATN GVHVHIDRVL TQIGTSIQDF 240
IEAEDDLSSF RAAAITSDIL EALGRDGHFT LFAPTNEAFE KLPRGVLERF MGDKVASEAL 300
MKYHILNTLQ CSESIMGGAV FETLEGNTIE IGCDDGSITV NGIKMVNKKD IVTNNGVIHL 360
IDQVLIPDSA KQVIELAGKQ QTTFTDLVAQ LGLASALRPD GEYTLAPVN NAFSDDTLMS 420

```

VQRLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKGSKQG 480
 RNGAIHIFRE IIKPAEKSLH EKLKQDKRFS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA 540
 FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKGFEPG VTNILKTQGG SKIFLKEVND 600
 TLLVNLKSK ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST 660
 FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIIKTEGPTL TKVKIEGEPE FRLIKEGETI 720
 TEVIHGEPII KKYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL 780
 QEEVTKVTKF IEGGDGHLFE DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ

BCA7 DNA sequence (SEQ ID NO:3)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset
 Accession #: Z29083; Nucleic Acid Accession #: NM_006670; Coding sequence: 85-1347 (start
 and stop codons underlined)

CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCAGCCC 60
 AGCTCCGGGG AAACGCGAGC CGCGATGCCT GGGGGGTGCT CCCGGGGGCC CGCCGCCGGG 120
 GACGGGCGTC TCGCGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCCTCGTCT 180
 TCTCCACCTT CCTCGGCATC CTCCTTCTCC TCCTCGGCGC CGTTCCTGGC TTCCGCCGTG 240
 TCCGCCCAGC CCCCGCTGCC GGACCACTGC CCCGCGCTGT GCGAGTGCTC CGAGGCAGCG 300
 CGCACAGTCA AGTGCCTTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC 360
 GTGCGCAACC TCTTCTTTAC CGGCAACCAG CTGGCCGTGC TCCCTGCCGG CGCCTTCGCC 420
 CGCCGCCCGC CGCTGGCGGA GCTGGCCGCG CTCAACCTCA GCGGCAGCCG CCTGGACGAG 480
 GTGCGCGCGG GCGCCTTCGA GCATCTGCCC AGCCTGCGCC AGCTCGACCT CAGCCACAAC 540
 CCATCGGCCG ACCTCAGTCC CTTCGCTTTC TCGGGCAGCA ATGCCAGCGT CTCGGCCCCC 600
 AGTCCCTTGT TGGAACTGAT CCTGAACCAC ATCGTGCCCC CTGAAGATGA GCGGCAGAAC 660
 CGGAGCTTCG AGGGCATGGT GGTGGCGGCG CTGCTGGCGG GCCGTGCACT GCAGGGGGCTC 720
 CGCCGCTTGG AGCTGGCCAG CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCAA 780
 CTGCCAGGCC TCAGGCACCT GGACTTAAGT AATAATTGCG TGGTGAGCCT GACCTACGTG 840
 TCCTTCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900
 CTTCACAATG GCACCTTGGC TGAGTTGCAA GGTCTACCCC ACATTAGGCT TTTCTTGAC 960
 AACAAATCCCT GGGTCTGCGA CTGCCACATG GCAGACATGG TGACCTGGCT CAAGGAAACA 1020
 GAGGTAGTGC AGGGCAAAGA CCGGCTCACC TGTGCATATC CGGAAAAAAT GAGGAATCGG 1080
 GTCTCTTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTCC CCCATCCCTG 1140
 CAAACCTCTT ATGTCTTCTT GGTATTGTT TTAGCCCTGA TAGGCGCTAT TTTCTCCTG 1200
 GTTTTGATTT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCCTGC 1260
 AGGGATCACA TGAAGGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA 1320
 AACCTCAGTT CTAACCTCGA TGCTGAGAA ATATTAGAGG ACAGACCAAG GACAACTCTG 1380
 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCCACTTTCA TCCTCCACTA 1440
 TAGATACAAC GGACTTTGAC TAAAGCAGT GAAGGGGATT TGCTTCCTTG TTATGTAAAG 1500
 TTTCTCGGTG TGTCTGTGTA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCCTC 1560
 TTCTTTTCTT TGAACCTCCT CAACACGTAT GGAGGGATT TTCAGGTTTC AGCATGAACA 1620
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680
 ACAGATAGCA TTAACAAAAA GCTGCCTCAA CTTTTCGAG AAAAATACTT TATTCAATAA 1740
 TATCAGTTT ATTCTCATGT ACCTAAGTTG TGGAGAAAAT AATTGCATCC TATAAACTGC 1800
 CTGCAGACGT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA 1860
 AGAGCATGCT TACATTTTAC TGTCTGTCAT ATTACAAAAA ATAACCTGCA ACTTCATAAC 1920
 TTCTTTGACA AAGTAAATTA CTTTTTGTAT TGCAGTTTAT ATGAAAATGT ACTGATTTT 1980
 TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAAATAA AAAAATAAAG 2040
 ATTCTTAAAA GAA

BCA7 Protein sequence (SEQ ID NO:4)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset
 Accession #: Z29083; Protein Accession #: NP_006661; Predicted Signal sequence: 1-32;
 Predicted TM domains: 357-373; PFAM domains: leucine_rich_repeats: 61-90, 119-142, 143-166,
 235-258, 259-282, 294-345;
 Summary: a type 1a TM protein of unknown function, detected in multiple cancers, with highest
 expression in breast cancer.

MPGGCSRGP AAGDRLRLAR LALVLLGWVS SSSPTSSASS FSSSAPFLAS AVSAQPPLPD 60
 QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTQ NQLAVLPAGA FARRPPLAEL 120
 AALNLSGSRL DEVRAFAFEH LPSLRQLDLS HNPLADLSPF AFSGSNASVS APSPLVELIL 180
 NHIVPPEDER QNRSFEGMVV AALLAGRALQ GLRRLELASN HFLYLPDVL AQLPSLRHLD 240
 LSNNLSVSLT YVSFRNLTHL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVF LDNNPWVDC 300
 HMADMVTLWK ETEVQVKDR LTCAYPEKMR NRVLLLENSA DLDCDPLPP SLQTSYVFLG 360
 IVLALIGAIF LVLVLYNRKG IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSSNSDV

BCX5 DNA sequence (SEQ ID NO:5)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Nucleic Acid
 Accession #: AF160477; Coding sequence: 225-1757 (start and stop codons underlined)

GGGGAGCTCG GAGCTCCCGA TCACGGCTTC TTGGGGGTAG CTACGGCTGG GTGTGTAGAA 60
 CGGGGCCCGG GCTGGGGCTG GGTCCCCTAG TGAGACCCAA GTGCGAGAGG CAAGAAGTCT 120
 GCAGCTTCCT GCCTTCTGGG TCAGTTCTT ATTCAAGTCT GCAGCCGGCT CCCAGGGAGA 180
 TCTCGGTGGA ACTTCAGAAA CGCTGGGCGA TCTGCCTTTC AACCATGCCC CTGTCCCTGG 240
 GAGCCGAGAT GTGGGGGCTT GAGGCCTGGC TGCTGCTGCT GCTACTGCTG GCATCATTTA 300
 CAGGCCGCTG CCCCGCGGGT GAGCTGGAGA CCTCAGACGT GGTAACCTGT GTGCTGGGCC 360
 AGGACGCAAA ACTGCCCTGC TTCTACCGAG GGGACTCCGG CGAGCAAGTG GGGCAAGTGG 420

CATGGGCTCG	GGTGGACGCG	GGCGAAGGCG	CCCAGGAACT	AGCGCTACTG	CACTCCAAAT	480
ACGGGCTTCA	TGTGAGCCCG	GCTTACGAGG	GCCGCGTGGA	GCAGCCGCGC	CCCCACGCA	540
ACCCCTGGA	CGGCTCAGTG	CTCCTGCGCA	ACGCAGTGCA	GGCGGATGAG	GGCGAGTACG	600
AGTGCCGGGT	CAGCACCTTC	CCCGCCGGCA	GCTTCCAGGC	GCGGCTGCGG	CTCCGAGTGA	660
TGGTGCCTCC	CCTGCCCTCA	CTGAATCCTG	GTCCAGCACT	AGAAGAGGGC	CAGGGCCTGA	720
CCCTGGCAGC	CTCCTGCACA	GCTGAGGGCA	GCCCAGCCCC	CAGCGTGACC	TGGGACACGG	780
AGGTCAAAGG	CACAACGTCC	AGCCGTTCCT	TCAAGCACTC	CCGCTCTGCT	GCCGTCACCT	840
CAGAGTTCCA	CTTGGTGCCT	AGCCGCAGCA	TGAATGGGCA	GCCACTGACT	TGTGTGGTGT	900
CCCATCCTGG	CCTGCTCCAG	GACCAAAGGA	TCACCCACAT	CCTCCACGTG	TCCTTCCTTG	960
CTGAGGCCTC	TGTGAGGGGC	CTTGAAGACC	AAAATCTGTG	GCACATTGGC	AGAGAAGGAG	1020
CTATGCTCAA	GTGCCTGAGT	GAAGGGCAGC	CCCCTCCCTC	ATACAACTGG	ACACGGCTGG	1080
ATGGGCTCT	GCCAGTGGG	GTACGAGTGG	ATGGGGACAC	TTTGGGCTTT	CCCCACTGA	1140
CCACTGAGCA	CAGCGGCATC	TACGCTTGCC	ATGTCAGCAA	TGAGTTCTCC	TCAAGGGATT	1200
CTCAGGTCAC	TGTGGATGTT	CTTGACCCCC	AGGAAGACTC	TGGGAAGCAG	GTGGACCTAG	1260
TGTGAGCCTC	GGTGGTGGTG	GTGGGTGTGA	TCGCCGCACT	CTTGTCTGCG	CTTCTGGTGG	1320
TGGTGGTGGT	GCTCATGTCC	CGATACCATC	GGCGCAAGGC	CCAGCAGATG	ACCCAGAAAT	1380
ATGAGGAGGA	GCTGACCCTG	ACCAGGGAGA	ACTCCATCCG	GAGGCTGCAT	TCCCATCACA	1440
CGGACCCAG	GAGGCAGCCG	GAGGAGAGTG	TAGGGCTGAG	AGCCGAGGGC	CACCCGTGATA	1500
GTCTCAAGGA	CAACAGTAGC	TGCTCTGTGA	TGAGTGAAGA	GCCCAGGGGC	CGCAGTTACT	1560
CCACGCTGAC	CACGGTGAGG	GAGATAGAAA	CACAGACTGA	ACTGCTGTCT	CCAGGCTCTG	1620
GGCGGGCCGA	GGAGGAGGAA	GATCAGGATG	AAGGCATCAA	ACAGGCCATG	AACCATTTTG	1680
TTCAGGAGAA	TGGGACCCTA	CGGGCCAAGC	CCACGGGCAA	TGGCATCTAC	ATCAATGGGC	1740
GGGGACACCT	GGTCTGACCC	AGGCCTGCCT	CCCTTCCCTA	GGCCTGGCTC	CTTCTGTTGA	1800
CATGGGAGAT	TTTAGCTCAT	CTTGGGGGCC	TCCTTAAACA	CCCCCATTTT	TTGCGGAAGA	1860
TGCTCCCCAT	CCCACTGACT	GCTTGACCTT	TACCTCCAAC	CCTTCTGTTC	ATCGGGAGGG	1920
CTCCACCAAT	TGAGTCTCTC	CCACCATGCA	TGCAGGTGAC	TGTGTGTGTG	CATGTGTGCC	1980
TGTGTGAGTG	TTGACTGACT	GTGTGTGTGT	GGAGGGGTGA	CTGTCCGTGG	AGGGGTGACT	2040
GTGTCCGTGG	TGTGTATTAT	GCTGTGATAT	CAGAGTCAAG	TGAACGTGTG	TGTATGTGCC	2100
ACGGGATTTG	AGTGGTTGCG	TGGGCAACAC	TGTCAGGGT	TGGCGTGTGT	GTGATGTGGC	2160
TGTGTGTGAC	CTCTGCCTGA	AAAAGCAGGT	ATTTTCTCAG	ACCCAGAGC	AGTATTAAATG	2220
ATGCAGAGGT	TGGAGGAGAG	AGGTGGAGAC	TGTGGCTCAG	ACCCAGGTGT	GCGGGCATAG	2280
CTGGAGCTGG	AATCTGCCTC	CGGTGTGAGG	GAACCTGTCT	CCTACCACTT	CGGAGCCATG	2340
GGGGCAAGTG	TGAAGCAGCC	AGTCCCTGGG	TCAGCCAGAG	GCTTGAACCTG	TTACAGAAGC	2400
CCTCTGCCCT	CTGGTGGCCT	CTGGGCCTGC	TGCATGTACA	TATTTTCTGT	AAATATACAT	2460
GCGCCGGGAG	CTTCTTGCG	GAATACTGCT	CCGAATCACT	TTTAATTTT	TTCTTTTTTT	2520
TTTCTTGCCC	TTTCCATTAG	TTGTATTTT	TATTTATTTT	TATTTTATT	TTTTTTTAGA	2580
GATGGAGTCT	CACATGTTG	CTCAGGCTGG	CCTTGAACCTC	CTGGGCTCAA	GCAATCCTCC	2640
TGCCTCAGCC	TCCCTAGTAG	CTGGGACTTT	AAGTGACAC	CACTGTGCCT	GCTTTGAATC	2700
CTTTACGAAG	AGAAAAA	AATTAAGAA	AGCCTTTAGA	TTTATCCAAT	GTTTACTACT	2760
GGGATTGCTT	AAAGTGAGGC	CCCTCCAACA	CCAGGGGGTT	AATTCCTGTG	ATTGTGAAAG	2820
GGGCTACTTC	CAAGGCATCT	TCATGCAGGC	AGCCCCCTGG	GAGGGCACCT	GAGAGCTGGT	2880
AGAGTCTGAA	ATTAGGGATG	TGAGCCTCGT	GGTACTAGAG	TAAGGTAAAA	TTGCATCCAC	2940
CATTGTTTGT	GATACCTTAG	GGAATTGCTT	GGACCTGGTG	ACAAGGGCTC	CTGTTCAATA	3000
GTGGTGTGG	TGAGAGAGAG	AGCAGTGATT	ATAGACCGAG	AGAGTAGGAG	TTGAGGTGAG	3060
GTGAAGGAGG	TGCTGGGGGT	GAGAATGTCG	CCTTTCCCC	TGGGTTTTGG	ATCACTAATT	3120
CAAGGCTCTT	CTGGATGTTT	CTCTGGGTG	GGGCTGGAGT	TCAATGAGGT	TTATTTTATG	3180
CTGGCCACC	CAGATACACT	CAGCCAGAAT	ACCTAGATT	AGTACCCCAA	CTCTCTTAG	3240
TCTGAAATCT	GCTGGATTTC	TGGCCTAAGG	GAGAGGCTCC	CATCCTTCGT	TCCCCAGCCA	3300
GCCTAGGACT	TCGAATGTGG	AGCCTGAAGA	TCTAAGATCC	TAACATGTAC	ATTTTATGTA	3360
AAATATGTGA	TATTTGTACA	TAAATGATA	TTCTGTTTTT	AAATAAACAG	ACAAAACCTG	3420
TTCAAAAAA	AAAAA	AAAAA				

BCX5 Protein sequence (SEQ ID NO:6)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Protein Accession #: AF160477; Predicted Signal sequence: 1-26; Predicted TM domains: 355-371; PFAM domains: IgSF_domain: 45-129, 162-225, 263-317; Summary: A type Ia TM protein; is a member of the immunoglobulin superfamily.

MPLSLGAEMW	GPEAWLLLLL	LLASFTGRCP	AGELETSADV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAVARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRV	TFPAGSFQAR	LRLRMVPPPL	PSLNPGPAL	EGQGLTLAAS	CTAEGSPAPS	180
VTWDTVEVKG	TSSRSFKHSR	SAAVTSEFHL	VPSRSMNGQP	LTCVVSHPLG	LQDQRITHIL	240
HVSFLAEASV	RGLEDQNLWH	IGREGAMLKC	LSEGQPPPSY	NWTRLDGPLP	SGVRVDGDTL	300
GFPPLTTEHS	GIYVCHVSNE	FSSRDSQVTV	DVLDPQEDSG	KQVDLVASV	VVVGVIALL	360
FCLLVVVVVL	MSRYHRRKAQ	QMTQKYEEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
EGHPDSLKDN	SSCSVMSEEP	EGRSYSTLTT	VREIETQTEL	LSPGSGRAEE	EEDQDEGIKQ	480
AMNHVQENG	TLRAKPTGNG	IYINGRGHLV				

mouse BCX5 Protein sequence (SEQ ID NO:7)

Gene name: mouse_LNIR; Unigene number: n/a; Probeset Accession #: BF168327; Protein Accession #: n/a; Predicted Signal sequence: 1-27; Predicted TM domains: 346-362; PFAM domains: IgSF_domains:44-126,166-221,259-313; Summary: This is the mouse orthologue of human BCX5; it is a type Ia TM protein of unknown function.

MPLSLGAEMW	GPEAWLRLLF	LASFTGQYSA	GELETSDVVT	VVLGQDAKL	CFYRGDPDEQ	60
VQQVAVARVD	PNEYXPGAGL	LHSKYGLHVN	PAYEDRVEQX	XHETFRRSVL	LRNAVQADEG	120

EYECRVSTFP SGSFQARMRL RVLVPPPLPSL NPGPPLEEGQ ADVAASCTAE GSPAPSVTWD 180
 TEVKGQTQSSR SFTHPRSAAV TSEFHLVPSR SMNGQPLTCV VSHPGLLQDR RITHTLQVAF 240
 LAEASVRGLE DQNLWQVGRE GATLKCLSEG QPPPKNWTR LDGPLPSGVR VKGDTLGFPP 300
 LTTEHSGVYX CHVSNELSSR DSQVTVEVLD PEDPGKQVDL VSASVIIVGV IAALLFCLLV 360
 VVVVLMsRYH RRKAQMQTK YEEELTLTRE NSIRRLHSHH SDPRSQPEES VGLRAEGHPD 420
 SLKDNSSCSV MSEEPEGRSY STLTTVREIE TQTELLSPGS GRTEEDDDQD EGIKQAMNHL 480
 CRKMGF

BCZ6 DNA sequence (SEQ ID NO:8)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Nucleic Acid Accession #: NM_002184;
 Coding sequence: 256-3012 (start and stop codons underlined)

GAGCAGCCAA AAGGCCCGCG GAGTCGCGCT GGGCCGCCCC GGCGCAGCTG AACCGGGGGC 60
 CGCGCCTGCC AGGCCGACGG GTCTGGCCCA GCCTGGCGCC AAGGGGTTTC TCGCTGTGG 120
 AGACGCGGAG GGTTCGAGCG GCGCGGCTTG AGTGAAACCC AATGGAACCA GCATGACATT 180
 TAGAAGTAGA AGACTTAGCT TCAATCCCT ACTCCTTCAC TACTAATTT TGTGATTGG 240
 AAATATCCGC GCAAGATGTT GACGTGTCAG ACTTGGGTAG TGCAAGCCTT GTTTATTTTC 300
 CTCACCACTG AATCTACAGG TGAACCTCTA GATCCATGTG GTTATATCAG TCCTGAATCT 360
 CCAGTTGTAC AACTTCATTC TAATTTCACT GCAGTTTGTG TGCTAAAGGA AAAATGTATG 420
 GATTATTTTC ATGTAAATGC TAATTACATT GTCTGGAAAA CAAACCATTT TACTATTCTT 480
 AAGGAGCAAT ATACTATCAT AAACAGAACA GCATCCAGTG TCACCTTTAC AGATATAGCT 540
 TCATTAAATA TTCAGCTCAC TTGCAACATT CTTACATTTC GACAGCTTGA ACAGAATGTT 600
 TATGGAATCA CAATAATTTT AGGCTTGCCT CCAGAAAAAC CTAAAAATTT GAGTTGCATT 660
 GTGAACGAGG GGAAGAAAAAT GAGGTGTGAG TGGGATGGTG GAAGGGAAAC ACACTTGGAG 720
 ACAAACTTCA CTTTAAATC CACTGTGTGAT TATTCTACTG TGTATTTTGT CAACATTGAA 780
 CGTGACACCC CCACCTCATG CACTGTGTGAT TATTCTACTG TGTATTTTGT CAACATTGAA 840
 GTCTGGGTAG AAGCAGAGAA TGCCCTTGGG AAGGTTACAT CAGATCATAT CAATTTTGAT 900
 CCTGTATATA AAGTGAAGCC CAATCCGCCA CATAATTTAT CAGTGATCAA CTCAGAGGAA 960
 CTGTCTAGTA TCTTAAATTT GACATGGACC AACCCAAAGTA TTAAGAGTGT TATAATACTA 1020
 AAATATAACA TTCAATATAG GACCAAAGAT GCCTCAACTT GGAGCCAGAT TCCTCCTGAA 1080
 GACACAGCAT CCACCCGATC TTCATTCACT GTCCAAGACC TTAACCTTTT TACAGAATAT 1140
 GTGTTTAGGA TTCGCTGTAT GAAGGAAGAT GGTAAAGGAT ACTGGAGTGA CTGGAGTGAA 1200
 GAAGCAAGTG GGATCACCTA TGAAGATAGA CCATCTAAAG CACCAAGTTT CTGGTATAAA 1260
 ATAGATCCAT CCCATACCTA AGGCTACAGA ACTGTACAAC TCGTGTGGAA GACATTGCCT 1320
 CCTTTTGAAG CCAATGGAAG AATCTTGGAT TATGAAGTGA CTCTCACAAG ATGGAATCA 1380
 CATTTACAAA ATTACACAGT TAATGCCACA AAACCTGACG TAAATCTCAC AAATGATCGC 1440
 TATCTAGCAA CCCTAACAGT AAGAAATCTT GTTGGCAAAT CAGATGCAGC TGTTTTAACT 1500
 ATCCCTGCCT GTGACTTTCA AGCTACTCAC CCTGTAATGG ATCTTAAAGC ATTCCCCAAA 1560
 GATAACATGC TTTGGGTGGA ATGGACTACT CCAAGGGAAT CTGTAAAGAA ATATATACTT 1620
 GAGTGGTGTG TGTATCAGA TAAAGCACCC TGTATCAGC ACTGGCAACA AGAAGATGGT 1680
 ACCGTGCATC GCACCTATTT AAGAGGGAAC TTAGCAGAGA GCAATGCTA TTTGATAACA 1740
 GTTACTCCAG TATATGCTGA TGGACCAGGA AGCCCTGAAT CCATAAAGGC ATACCTTAAA 1800
 CAAGCTCCAC CTTCCAAAGG ACCTACTGTT CGGACAAAAA AAGTAGGGAA AACGGAAGCT 1860
 GTCTTAGAGT GGGACCAACT TCCTGTTGAT GTTCAGAATG GATTATATCAG AAATATACT 1920
 ATATTTTATA GAACCATCAT TGGAAATGAA ACTGCTGTGA ATGTGGATTG TTCCACACA 1980
 GAATATACAT TGTCTCTTT GACTAGTGAC ACATGTGACA TGGTACGAAT GGCAGCATAC 2040
 ACAGATGAAG GTGGGAAGGA TGGTCCAGAA TTCACCTTTA CTACCCCAA GTTTGCTCAA 2100
 GGAGAAATG AAGCCATAGT CGTGCTGTG TGTCTAGCAT TCCTATTGAC AACTCTTCTG 2160
 GGAGTGTCTG TCTGCTTTAA TAAGCGAGAC CTAATTAATA AACACATCTG GCCTAATGTT 2220
 CCAGATCCCT CAAAGAGTCA TATTGCCAG TGGTCACCTC ACACCTCTCC AAGGCACAAT 2280
 TTTAATTCAA AAGATCAAT GTATTCAGAT GGCAATTTCA CTGATGTAAG TGTGTGGAA 2340
 ATAGAAGCAA ATGACAAAAA GCCTTTTCCA GAAGATCTGA AATCATTGGA CCTGTTCAA 2400
 AAGGAAAAAA TTAATACTGA AGGACACAGC AGTGGTATTG GGGGGTCTTC ATGCATGTCA 2460
 TCTTCTAGGC CAAGCATTTC TAGCAGTGAT GAAAATGAAT CTTACAAAA CACTTCGAGC 2520
 ACTGTCCAGT ATTCTACCGT GGTACACAGT GGCTACAGAC ACCAAGTTCC GTCAGTCCAA 2580
 GTCTTCTCAA GATCCGAGTC TACCCAGCCC TTGTTAGATT CAGAGGAGCG GCCAGAAGAT 2640
 CTACAATTAG TAGATCATGT AGATGGCGGT GATGGTATTT TGCCCAGGCA ACAGTACTTC 2700
 AAACAGAACT GCAGTCAGCA TGAATCCAGT CCAGATATTT CACATTTTGA AAGGTCAAAG 2760
 CAAGTTTCAT CAGTCAATGA GGAAGATTTT GTTAGACTTA AACAGCAGAT TTCAGATCAT 2820
 ATTTTCAAT CCTGTGGATC TGGGCAAATG AAAATGTTTC AGGAAGTTTC TGCAGCAGAT 2880
 GCTTTTGGTC CAGGTACTGA GGGACAAGTA GAAAGATTTG AAACAGTTGG CATGGAGGCT 2940
 GCGACTGATG AAGGCATGCC TAAAGTTTAC TTACCACAGA CTGTACGCA AGGCGGCTAC 3000
 ATGCCTCAGT GAAGGACTAG TAGTTCCTGC TACAACCTCA GCAGTACCTA TAAAGTAAAG 3060
 CTAATAATGAT TTTATCTGTG AATTC

BCZ6 Protein sequence (SEQ ID NO:9)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Protein Accession #: NP_002175; Predicted
 Signal sequence: 1-22; Predicted TM domains: 625-641; PFAM domains:
 fibronectin_type_III_domains: 222-311, 424-509, 519-606; Summary: A type I TM protein; it
 homodimerizes or heterodimerized to make a functional receptor for IL-6, oncostatin-M, IL-11,
 LIF, and CNTF.

MLTLQTWVQV ALFIFLTES TGEILLPCGY ISPESPVVQL HSNFTAVCVL KEKCMDYFHV 60
 NANYIVWKTN HFTIPKEQYT IINRTASSVT FTDIASLNIQ LTCNLTFGQ LEQNVYGITI 120

ISGLPPEKPK	NLSCIVNEGK	KMRCEWDGGR	ETHLETNFTL	KSEWATHKFA	DKAKRDTPT	180
SCTVDYSTVY	FVNIEVWVEA	ENALGKVTSD	HINFDPVYKV	KPNPPHNLVS	INSEELSSIL	240
KLTWTNPSIK	SVIILKYNIQ	YRTKDASTWS	QIPPEDTAST	RSSFVQDLK	PFTEYVFRIR	300
CMKEDGKGYW	SDWSEEASGI	TYEDRPSKAP	SFWYKIDPSH	TQGYRTVQLV	WKTLPPEAN	360
GKILDYEVTL	TRWKSHLQNY	TVNATKLTVN	LTNDRYLATL	TVRNLVGKSD	AAVLTIPACD	420
FQATHPVMDL	KAFPKDNMLW	VEWTTPRESV	KKYILEWCVL	SDKAPCITDW	QQEDGTVHRT	480
YLRGNLAESK	CYLITVTPVY	ADGPGSPESI	KAYLKQAPPS	KGPTVRTKKV	GKNEAVLEWD	540
QLPVDVQNGF	IRNYTIFYRT	IIGNETAVNV	DSSHTEYTLS	SLTSDTLYMV	RMAAYTDEGG	600
KDGPEFTFTT	PKFAQGEIEA	IVVPVCLAFI	LTTLLGVLCF	FNKRDLIKKH	IWPNVPDPSK	660
SHIAQWSPHT	PPRHNFNISK	QMYSDGNFTD	VSVVEIEAND	KKPFPEDLKS	LDLFFKKEKIN	720
TEGHSSGIGG	SSCMSSSRPS	ISSSDENESS	QNTSSTVQYS	TVVHSGYRHQ	VPSVQVFSRS	780
ESTOPLDSE	ERPEDLQLVD	HVDGGDGILP	RQQYFKQNC	QHSSPDISH	FERSKQVSSV	840
NEEDFVRLKQ	QISDHISQSC	GSGQMCMFQE	VSAADAFGPG	TEGQVERFET	VGMEAATDEG	900
MPKSYLPQTV	RQGGYMPQ					

BFG4 DNA sequence (SEQ ID NO:10)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
Nucleic Acid Accession #: AB020689; Coding sequence: 108-2777 (start and stop codons underlined)

GAACCTATGT	AGCCTCATTA	TCCCGCTCCG	TGAGGTGACA	ATTGTGGAAA	AGGCAGACAG	60
CTCCAGTGTG	CTCCCCAGTC	CCTTATCACA	TCAGCACCCG	AAACAGGATG	ACCTTCCTAT	120
TTGCCAACTT	GAAAGATAGA	GACTTTCTAG	TGCAGAGGAT	CTCAGATTTC	CTGCAACAGA	180
CTACTTCCAA	AATATATTCT	GACAAGGAGT	TTGCAGGAAG	TTACAACAGT	TCAGATGATG	240
AGGTGTACTC	TCGACCACAG	AGCCTCGTCT	CCTCCAGCCC	CCAGAGAAGC	ACGAGCTCTG	300
ATGCTGATGG	AGAGCGCCAG	TTTAACCTAA	ATGGCAACAG	CGTCCCCACA	GCCACACAGA	360
CCCTGATGAC	CATGTATCGG	CGGCGGTCTC	CCGAGGAGTT	CAACCCGAAA	TTGGCCAAAG	420
AGTTTCTGAA	AGAGCAAGCC	TGGAAGATTG	ACTTTGTCTG	GTATGGGCAA	GGGATCTGCA	480
TGTACCGCAC	AGAGAAAACG	CGGGAGCTGG	TGTTGAAGGG	CATCCCGGAG	AGCATGCGTG	540
GGGAGCTCTG	GCTGCTGCTG	TCAGGTGCCA	TCAATGAGAA	GGCCACACAT	CCTGGGTACT	600
ATGAAGACCT	AGTGGAGAAG	TCCATGGGGA	AGTATAATCT	CGCCACGGAG	GAGATTGAGA	660
GGGATTTACA	CGCTCCCTTT	CCAGAACACC	CAGCTTTTCA	GAATGAAATG	GGCATTGCTG	720
CACTAAGGAG	AGTCTTAAAC	GCTTATGCTT	TTCGAAATCC	CAACATAGGG	TATTGCCAGG	780
CCATGAATAT	TGTCACCTCA	GTGCTGTGCT	TTTATGCCAA	AGAGGAGGAA	GCTTTCTGGC	840
TGCTTGTGGC	TTTGTGTGAG	CGCATGCTCC	CAGATTACTA	CAACACCAGA	GTTGTGGGTG	900
CACCTGGTGA	CCAAGGTGTC	TTTGAGGAGC	TAGCACGAGA	CTACGTCCCA	CAGCTGTACG	960
ACTGCATGCA	AGACCTGGGC	GTGATTTCCA	CCATCTCCCT	GTCTTGSTTC	CTCACACTAT	1020
TTCTCAGTGT	GATGCCTTTT	GAGAGTGCAG	TTGTGGTTGT	TGACTGTTTC	TTCTATGAAG	1080
GAATTAAGT	GATATTCCAG	TTGGCCCTAG	CTGTGCTGGA	TGCAAAATGT	GACAAACTGT	1140
TGAACCTCAA	GGATGATGGG	GAGGCCATGA	CCGTTTTGGG	AAGGTATTTA	GACAGTGTGA	1200
CCAATAAAGA	CAGCACACTG	CCTCCCATTG	CTCACCTCCA	CTCCTTGCTC	AGCGATGATG	1260
TGGAACCTTA	CCCTGAGGTA	GACATCTTTA	GACTCATCAG	AACTTCCTAC	GAGAAATTCC	1320
GAACATCCG	GGCAGATTGT	ATTGAACAGA	TGAGATTCAA	ACAGAGACTG	AAAGTGATCC	1380
AGACGCTGGA	GGTACTACG	AAACGCAACG	TGGTACGAAC	CATTGTGACA	GAAACTTCCT	1440
TTACCATTGA	TGAGCTGGAA	GAACTTTATG	CTCTTTTCAA	GGCAGAACAT	CTCACCAGCT	1500
GCTACTGGGG	CGGGAGCAGC	AACGCGCTGG	ACCGGCATGA	CCCCAGCCTG	CCCTACCTGG	1560
AACAGTATCG	CATTGACTTC	GAGCAGTTCA	AGGGAATGTT	TGCTCTTCTC	TTTCCCTGGG	1620
CATGTGGAA	TCACCTGAC	GTTCTGGCCT	CCCGCTTGTT	CCAGTTATTA	GATGAAAATG	1680
GAGACTCTTT	GATTAACTTC	CGGGAGTTTG	TCTCTGGGCT	AAGTGCTGCA	TGCCATGGGG	1740
ACCTCACAGA	GAAGCTCAAA	CTCCTGTACA	AAATGCACGT	CTTGCTGAG	CCATCCTCTG	1800
ATCAAGATGA	ACCAAGTTCT	GCTTTTGAAG	CAACTCAGTA	CTTCTTTGAA	GATATTACCC	1860
CAGAATGTAC	ACATGTTGTT	GGATTGGATA	GCAGAAGCAA	ACAGGGTGCA	GATGATGGCT	1920
TTGTTACGGT	GAGCCTAAAG	CCAGACAAAG	GGAAGAGAGC	AAATTCCTCA	GAAAATCGTA	1980
ATTATTTGAG	ACTGTGGACT	CCAGAAAATA	AATCTAAGTC	AAAGAATGCA	AAGGATTTAC	2040
CCAAATTAAA	TCAGGGGCAG	TTCAATTGAAC	TGTGTAAGAC	AATGTATAAC	ATGTTACAGC	2100
AAGACCCCAA	TGAGCAGGAG	CTGTACCATG	CCACGGCAGC	AGTGACCAGC	CTCCTGCTGG	2160
AGATTGGGGA	GGTCCGCAAG	TTGTTTCGTG	CCCAGCCTGC	AAAGGAGGGC	GGGAGCGGAG	2220
GCACTGGGCC	GTCTTGCCAC	CAGGGCATCC	CAGGCGTGCT	CTTCCCCAAG	AAAGGGCCAG	2280
GCCAGCCTTA	CGTGGTGGAG	TCTGTTGAGC	CCCTGCGGCG	CAGCCTGGCC	CCCGACAGCG	2340
AGGAACACTC	CCTTGGAGGA	CAATGGAGG	ACATCAAGCT	GGAGGACTCC	TCGCCCCGGG	2400
ACAACGGGGC	CTGCTCTCTC	ATGCTGATCT	CTGACGACGA	CACCAAGGAC	GACAGCTCCA	2460
TGTCCTCATA	CTCGGTGCTG	AGTGCCGGCT	CCCACGAGGA	GGACAAGCTG	CACTGCGAGG	2520
AAATCGGAGA	GGACACGGTC	CTGGTGCGGA	GCGGCCAGGG	CACGGCGGCA	CTGCCCCGGA	2580
GCACCAGCCT	GGACCGGGAC	TGGGCCATCA	CCTTCGAGCA	GTTCTTGCC	TCCCTCTTAA	2640
CTGAGCCTGC	CCTGCTCAAG	TACTTTGACA	AGCCCGTGTG	CATGATGGCC	AGGATTACCA	2700
GTGCAAAAAA	CATCCGGATG	ATGGGCAAGC	CCCTCACCTC	GGCCAGTGAC	TATGAAATCT	2760
CGGCCATGTC	CGGCTGACAC	GGGCGCCTTC	CCGGGGGAGT	GGGAGGAGAG	GGAGGGGAGG	2820
GATTTTTTAT	GTTCTTCTGT	GTTGAGTTTT	TCTTTTCTTT	CTTTTAAATT	AAATATTAT	2880
TAGTACCTGG	TATTGAAGCC	TAGTGTTTTC	ATAATGTAAT	TCAATGAAAA	CTGTTGGAGA	2940
AATATTAAAA	CACCTCAATG	TAGGTACATT	ACACTCTTGT	TGCGGGGAGG	GGATTTACCA	3000
GAATACAGTT	TATTTCTGTA	ATTCTAAAAA	ACAAAAAGAT	GAATCTGTCA	GTGATATGTG	3060
TGTATTATAA	TATTATTAAT	TTGCTGTGTA	GCTGTATACA	TGGTTTAAAA	AATAGTACTG	3120
TTTAATGCTA	AGTAAGGCAG	CAGTCATTGG	TGTATTCAAG	CTTTTTAAAT	AAAATTAGAG	3180
CTGTAAGGAA	AATGAAAGC	CACAAATGCA	AGACTGTTCT	TAAATGGAAG	GCATAGTCAG	3240
CGAGGGTAAA	TCTATACCA	CTTTAGGAAG	TATTAATAAT	ATTTTAAAGA	TTTGAAATAT	3300
ATTTTCATAGA	AGTCCTCTAT	TCAAAATCAT	ATTCCACAGA	TGTTCCCTTT	CAAAGGGAAA	3360

ACATTGTTGGG TTCTAAACAG TTATGAAAGT AAGTGATTTT TACATGATTC CAGAATAACA 3420
 CTGTGATTGA CCAATTTAGA CAGATACCCAG ACCAATTTTG CATTTAAGAA ATTGTTCTGA 3480
 TTATTTACGT CAACTCATT GAATTCAGTG AAAAGTAACA GTCTTTTGTC ACAGAGAATC 3540
 TGAAAGTAGC AGCAAAGACA GAGGGCTCAT GACAGGTTTT TGCTTTTGCT TTGCTTTGT 3600
 TTTTGAAGA GTAAAGTAC TGATGCTTCT GATACTGGAT GTTTAGCTTC TTAGTGCAA 3660
 AACATAAGTA AACAGTCAA CTTTACCATT TCCGTATTCT CCATAGATTG AAGAAATTTA 3720
 TACCACATAT CGCATATGAC CATCTTTCCA TCAAATCAAT GTAGAGATAA TGTAAGTGA 3780
 AAAAAAATCT GCAAGATAAT GTAAGTGAAT GTTTTAAAA CAGAACTTGT CACTTTATAT 3840
 AAAAGAATAG TATGCTCTAT TTCCTGAATG GATGTGGAAA TGAAAGCTAG CGCACCTGCA 3900
 CTTTGAATTC TTGCTTCTTT TTTATTACTG TTATGATTTT GCTTTTACA GATGTGGAC 3960
 GATTTTCTCT TCTGATTGTT GAATTCATAA TCATGGTCTC ATTTCTTTTG CTTCTTTGA 4020
 ATATTTCTTT CAACACATTCT CTTTATTTTA TTATACATTG TGTCTTTT TTAGCTATTG 4080
 CTGCTGTTGT TTTTATTCT ATTTACAGGA TGATTTTAA ACTGTCAAAT GAAGTAGTGT 4140
 TAACCTCAA TAGGCTAAAT GTGAACAAAT AAAATACAGC AAATACTCAG AAAAAA 4200
 AAAAAA AAAAA

BFG4 Protein sequence (SEQ ID NO:11)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
 Protein Accession #: BAA74905; Signal sequence: none; Predicted TM domains: 302-318; PFAM
 domains: TBC_domain: 135-347; Summary: a Type II membrane protein, likely localized to the
 peroxisome.

MTFLFANLKD RDLVQRISD FLQQTTSKIY SDKEFAGSYN SSDDEVYSRP SSLVSSSPQR 60
 STSSDADGER QFNLNNSVP TATQTLMTY RRRSPEEFNP KLAKEFLKEQ AWKIHFAEYG 120
 QGICMYRTEK TRELVLKGIP ESMRGELWLL LSGAINEKAT HPGYYEDLVE KSMGKYNLAT 180
 EEIERDLHRS LPEHPAFQNE MGIAALRRVL TAYAFRNPNI GYCOAMNIPT SVLLLYAKEE 240
 EAFWLLVALC ERMPLDYYNT RVVVALVDQG VFEELARDYV POLYDCMQDL GVISTISLSW 300
 FLTLFLSVMP PESAVVVVDC PFYEGIKVIF QLALAVLDAN VDKLLNCKDD GEAMTVLGRY 360
 LDSVTNKDST LPPIPHLHSL LSDDVEPYPE VDIFRLIRTS YEKFGTIRAD LIEQMRFKQR 420
 LKVIQTLED TCRNVVRTIV TETSFTIDEL EELYALFKAE HLTSCYWGGG SNALDRHDP 480
 LPYLEQYRID FEQFKGMFAL LFPWACGTHS DVLASRLFQL LDENGDSLIN FREFVSGLSA 540
 ACHGDLTEKL KLLYKMHVLP EPSSDQDEPD SAFEATQYFF EDITPECTHV VGLDSRSKQG 600
 ADDGFVTVSL KPDKGKRANS QENRNYLRW TPENKSKSKN AKDLPLKLNQG QFIELCKTMY 660
 NMFSDEPNEQ ELYHATAAVT SLLLEIGEVG KLFVAQPAKE GSGSGSGPSC HQGIPGVLP 720
 KKGPQPYPVV ESVEPLPASL APDSEHSLG QMEDIKLED SSPRDNGACS SMLISDDDTK 780
 DSSMSYSV LSAGSHEEDK LHCEEIGEDT VLVRSGQGT ALPRSTSLDR DWAITFEQFL 840
 ASLLTEPALV KYFDKPVCM ARITSARNIR MMGKPLTSAS DYEISAMSG

BCU7 DNA sequence (SEQ ID NO:12)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Nucleic Acid
 Accession #: n/a; Coding sequence: 1-573 (stop codon underlined)

TATTTTATTT TCCAGGCTAA AGCAAATGAA AGTTTGCTGG TATCAACACA GCCTGCCATA 60
 TTTTTCACAG CATGCAACAA TGGTGCTAGG ATAGCTATTT CTTACTGTAA TTGCCAGAGG 120
 CAGAAATGGT CTGGGTATAA GCTATTTTCAT AAAAGCAGCT TTAAATTGTC AGTATTAAGG 180
 TTTTCATGTG GAAAGGTGTC ATTCAAAAAA AAAGTAATTG GCATACATAT TCCACATCAT 240
 CGATCCTCTC TGTGGTGTTA ATTTTATTTAT ATGACCAGTA GAAAAATTTT AATATTTCTCA 300
 CAATATAGGT TTTGGGGCTT CCATATCATC AAAAGACTGA AAAATTATAA TTTTAGAATT 360
 AAACGTATGG ATTTTCATTAT AGAATTATCT GTGAGTTGTG TAGACACAGT CTTAATGTTT 420
 CTGGTTATGA CAGATAAGTT TGCTCAAAAA ATGTGGATGA AGCCATTATT GTTATTATTG 480
 TTATTGCTTC TGTTTCAGTTG TCTAAGTATC ATCCCTTCTG TGGCCCATCA CGCAGCAGAG 540
 TTGCCCTACA AATTTTCATTT GGCAGCGCCA TAACATTCAT TTAATAAGTT TATGAAAAACA 600
 TTCAATTTGAA AGTTCCATGC AGCTTTAGCA CAGAGTTGAC CAAACACTGG CGTAAGTTCA 660
 ATTTACACAG AATATTGAA TTGAAACAAT AGAAATTTT CTCATAATAT ATACCTATGT 720
 GAAACCAACT TATCTGCATA ATTAAATCTA ATACATATTT AAGCCAGTTT AAGTGCTTTG 780
 TGTTGATGCC ATGCTTATCA AATACATGCA CAAGCTAAAC ATAATTGAA TGGGTCTATG 840
 AAGGAAAAAT AATGCTTAGA CTTTGGTGTA GGTCTCTCCT GTGTAGCCAT ATACCCAGGC 900
 TCTGCAGTAT CGAAGGATGC AAATGTTGAC ATAGATGGAA GCTCTTACCT ACCAAAGTGT 960
 TTAGGAAGGA TAAAGTTACA TTTGTCTTAA TTTCTAACAT TATCTTTGCT TTTATGTTTC 1020
 ATAAAAATTT GTCATTATTT ATGCTGGTGA AACGTATAAT CACATCCAAT TATTTGAACA 1080
 CATGCAAAAT AATTTTAAAT ATTATGTTAT TGTTTAAATT TGACTTATGG GAGATCAGTC 1140
 AAAAACTTAG AAGGTTTAACT ACTTCACTGA TTAATGGTGC TGAAAACACG TTACAATTAC 1200
 CACATATCCT TGCTATAAGT TTTGAAGTTT CTTAGCAATT AAAGTTTTTT TATTCAGTGT 1260
 GAACTGTGAG TATCTATTCT GGTGCTAAAT GTATGGTGCT AAATGAATTG TTAGTGTGA 1320
 TGGCTTAGT AATGCTCCTT TTATTCATTG CTAAATTTAG TGTATCCAT TTGATTCCTG 1380
 ATTCAGAAAT ATCAATAAAA TCCTATGTTA AATTAATCTT TACCAAAAAC AGGCAAGTTA 1440
 ACTCTGTTGT TTTAATTCAA CAGTCCAACA TTATTTAGGT GTTACAGAGT GTAAATATAT 1500
 TTCTTTGGGA GTTATTTCT TTTTAAAAAT TTTTATAGC TTGGCAATGT CCAAAGTCAA 1560
 ATATCACCTA AACTGGTTAG ATTACTTCTA CAGCTAATAA TATTGCAGGC ACTGGCGCCC 1620
 TCTGGTGGTT ATGAAGACAA ATTCTTAATG GCTACTTGAC CTACAGCAAA AGCCATTTCT 1680
 GTACCATAAA AATTTGTTGT GCAATATTAG AATTATCATA TGTTTCCTAC ATCTGACAGC 1740
 ACCTAAAAAT TTTGATAATA TTAACATGTA TCTAAGAGGA AAAAAAGATT AATATATTCT 1800
 GGCACCCACT TTCCTAGTAA TGTTTCCCAT GATTTCCAG TTCTGAGGCA CTTATTAAAG 1860
 TGCTTTTTTT TTTCTGAATT AATTAGGTAT TGGTAAAAA TATTTTTAAA TTTAGTTAGC 1920
 TTTATAACA CAATTAGAAT TACAATTAAT TAACAGAGGT ATAATTGTCT CACTTTCAGA 1980
 AGTGATCATT TATTTTATT TAGCACAGGT CATAAGAAAA ATATATAGAA AAATAATCAA 2040

```

TTTCATATAT AAAAGGATTA TTTCTCCACC TTTAATTATT GGCCTATCAT TTGTTAGTGT 2100
TATTTGGTCA TATTATTGAA CTAATGTATT ATTCCATTCA AAGTCTTTCT AGATTTAAAA 2160
ATGTATGCAA AAGCTTAGGA TTATATCATG TGTAACATT ATAGATAACA TCCTAAACCT 2220
TCAGTTTAGA TATATAATTG ACTGGGTGTA ATCTCTTTTG TAATCTGTTT TGACAGATTT 2280
CTTAAATTAT GTTAGCATAA TCAAGGAAGA TTTACCTTGA AGCACTTTCC AAATTGATAC 2340
TTTCAAACCT ATTTTAAAGC AGTAGAACCT TTTCTATGAA CTAAATCACA TGCAAAACTC 2400
CAACCTGTAG TATACATAAA ATGGACTTAC TTATTCCTCT CACCTTCTCC AGTGCCTAGG 2460
AATATTCTTC TCTGAGCCCT AGGATTGATT CTATCACACA GAGCAACATT AATCTAAATG 2520
GTTTAGCTCC CTCTTTTTC TCTAAAAACA ATCAGCTAAT AAAAAAAAAA TTTGAGGGCC 2580
TAAATTATTT CAATGGTTGT TTGAAATATT CAGTTCAGTT TGTACCTGTT AGCAGTCTTT 2640
CAGTTTGGGG GAGAATTAAT TACTGTGCTA AGCTGGTGCT TGGATACATA TTACAGCATC 2700
TTGTGTTTTA TTTGACAAAC AGAATTTTGG TGCCATAATA TTTTGAGAAT TAGAGAAGAT 2760
TGTGATGCAT ATATATAAAC ACTATTTTAA AAAAATATCT AAATATGTCT CACATATTTA 2820
TATAATCCTC AAATATACTG TACCATTTTA GATATTTTTT AAACAGATTA ATTTGGAGAA 2880
GTTTTATTCA TTACCTAATT CTGTGGCAA AATGGTGCT CTGATGTTGT GATATAGTAT 2940
TGTCAGTGTG TACATATATA AAACCTGTGT AAACCTCTGT CCTTATGAAC CATAACAAAT 3000
GTAGCTTTTT AAAGTCCATT GTATTGTTTT TTCTTTCAAT AAAAGAGTAT AATTAATTGG 3060
TTGTTTTTGA

```

BCU7 Protein sequence (SEQ ID NO:13)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 125-141, 154-170; PFAM domains: none; Summary: A type III membrane protein, highly overexpressed in breast cancer and prostate cancer; unknown function.

```

YFIFQAKANE SLLVSTQPAI FFTACNNGAR IAISYCNCQR QKWSGYKLFH KSSFKLSVLR 60
FSCGKVSFKK KVIGIHIPHH RSSLWCXFFY MTSRKILIFS QYRFWGFHII KRLKNYNFRI 120
KLMDFIIELS VSCVDTVLMF LVMTDKFAQK MWMKPLLLLL LLLLFSCLSI IPSVAHHAAE 180
LPYKFHLAAP

```

BFA1 DNA sequence (SEQ ID NO:14)

Gene name: calsynenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Nucleic Acid Accession #: NM_022131; Coding sequence: 11-2878 (start and stop codons underlined)

```

TGCTGCGAGG ATGCTGCCTG GCGGGCTGTG CTGGGTGCCG CTCCTGCTGG CGCTGGGCGT 60
GGGGAGCGGC AGCGGCGGTG GCGGGGACAG CCGGCAGCGC CGCTCCTCG CGGCTAAAGT 120
CAATAAGCAC AAGCCATGGA TCGAGACTTC ATATCATGGA GTCATACTG AGAACAATGA 180
CACAGTCATT TTGGACCCAC CACTGGTAGC CCTGGATAAA GATGCACCGG TTCTTTTTCG 240
AGGGGAAATC TGTGCGTTCA AGATCCATGG CCAGGAGCTG CCCTTTGAGG CTGTGGTGCT 300
CAACAAGACA TCAGGAGAGG GCCGGCTCCG TGCCAAGAGC CCCATTGACT GTGAGTTGCA 360
GAAGGAGTAC ACATTCATCA TCCAGGCCCTA TGAAGTGCTT GCTGGGCCCC ACGAGACAGC 420
CTGGAAAAAG TCACACAAGG CCGTGGTCCA TATACAGGTG AAGGATGTCA ACGAGTTTGC 480
TCCCACCTTC AAAGAGCCAG CCTACAAGGC TGTGTGACG GAGGGCAAGA TCTATGACAG 540
CATTCTGCAG GTGGAGGCCA TTGACGAGGA CTGCTCCCCA CAGTACAGCC AGATCTGCAA 600
CTATGAAATC GTCACCACAG ATGTGCCTTT TGCCATCGAC AGAAATGGCA ACATCAGGAA 660
CACTGAGAAG CTGAGCTATG ACAAACAACA CCAGTATGAG ATCCTGGTGA CCGCTACGA 720
CTGTGGACAG AAGCCCGCTG CTCAGGACAC CCGGTGTCAG GTGGATGTGA AGCCAGTTTG 780
CAAGCCTGGC TGGCAAGACT GGACCAAGAG GATTGAGTAC CAGCCTGGCT CCGGGAGCAT 840
GCCCTGTTC CCCAGCATCC ACCTGGAGAG GTGCGATGGA GCCGTGTCTT CCCTCCAGAT 900
CGTCACAGAG CTGCAGACTA ATTACATTGG GAAGGGTTGT GACCGGGAGA CCTACTCTGA 960
GAAATCCCTT CAGAAGTTAT GTGGAGCCTC CTCTGGCATC ATTGACCTCT TGCCATCCCC 1020
TAGCGCTGCC ACCAACTGGA CTGCAGGACT GCTGGTGGAC AGCAGTGAGA TGATCTTCAA 1080
GTTTGACGGC AGGCAGGGTG CCAAATCCC CGATGGGATT GTGCCAAGA ACCTGACCGA 1140
TCAGTTACCC ATCACCATGT GGATGAAACA CGGCCCCAGC CCTGGTGTGA GAGCCGAGAA 1200
GGAAACCATC CTCTGCAACT CAGACAAAAC CGAAATGAAC CGGCATCACT ATGCCCTGTA 1260
TGTGCACAAC TGCCGCTCTG TCTTTCTCTT GCGGAAGGAC TTCGACCAGG CTGACACCTT 1320
TCGCCCCGCG GAGTTCCACT GGAAGCTGGA TCAGATTTGT GACAAAGAGT GGCCTACTA 1380
TGTCAATCAAT GTGGAGTTTC CTGTGGTAAC CTTATACATG GATGGAGCAA CATATGAACC 1440
ATACCTGGTG ACCAACGACT GGCCCATTTA TCCATCTCAC ATAGCCATGC AACTCACAGT 1500
CGGCGCTTGT TGGCAAGGAG GAGAAGTCAC CAAACCACAG TTTGCTCAGT TCTTTCATGG 1560
AAGCCTGGCC AGTCTCACCA TCCGCCCTGG CAAAATGGAA AGCCAGAAGG TGATCTCCTG 1620
CCTGCAGGCC TGCAAGGAAG GGCTGGACAT TAATTCCTTG GAAAGCCTTG GCCAAGGAAT 1680
AAAGTATCAC TTCAACCCCT CGCAGTCCAT CCTGGTGATG GAAGGTGACG ACATTGGGAA 1740
CATTAACCGT GCTCTCCAGA AAGTCTCCTA CATCAACTCC AGGCAGTTCC CAACGGCGGG 1800
TGTGCGGGCC CTCAAAGTAT CCTCCAAAGT CCAGTGCTTT GGGGAAGACG TATGCATCAG 1860
TATCCCTGAG GTAGATGCCT ATGTGATGGT CCTCCAGGCC ATCGAGCCCC GGATCACCTT 1920
CCGGGGCACA GACCATTCTT GGAGACCTGC TGCCCAAGTTT GAAAGTGCCA GGGGAGTGAC 1980
CCTCTTCCCT GATATCAAGA TTGTGAGCAC CTTGCCCCAA ACCGAAGCCC CCGGGGACGT 2040
GAAAACCACA GACCCCAAT CAGAAGTCTT AGAGGAAATG CTTTATAACT TAGATTCTTG 2100
TGACATTTTG GTGATCGGAG GGGACTTGGG CCCAAGGCAG GAGTGCTTGG AGCTCAACCA 2160
CAGTGAGCTC CACCAACGAC ACCTGGATGC CACTAATTCT ACTGCAGGCT ACTCCATCTA 2220
CGGTGTGGGC TCCATGAGCC GCTATGAGCA GGTGCTACAT CACATCCGCT ACCGCAACTG 2280
GCGTCCGGCT TCCCTTGAGG CCCGGCGTTT CCGGATTAAG TGCTCAGAAC TCAATGGGCG 2340
TTACACTAGC AATGAGTTCA ACTTGGAGGT CAGCATCCTT CATGAAGACC AAGTCTCAGA 2400
TAAGGAGCAT GTCAATCATC TGATTGTGCA GCCTCCCTTC CTCCAGTCTG TCCATCATCC 2460
TGAGTCCCGG AGTAGCATCC AGCACAGTTC AGTGGTCCCA AGCATTGCCA CAGTGGTCAT 2520

```

CATCATCTCC GTGTGCATGC TTGTGTTTGT CGTGGCCATG GGTGTGTACC GGGTCCGGAT 2580
CGCCACCAG CACTTCATCC AGGAGACTGA GGCTGCCAAG GAATCTGAGA TGGACTGGGA 2640
CGATTCTGGG CTGACTATCA CAGTCAACCC CATGGAGAAA CATGAAGGAC CAGGGCATGG 2700
GGAAGATGAG ACTGAGGGAG AAGAGGAGGA AGAAGCCGAG GAAGAAATGA GCTCCAGCAG 2760
TGGCTCTGAC GACAGCGAAG AGGAGGAGGA GGAGGAAGGG ATGGGCAGAG GCAGACATGG 2820
GCAGAATGGA GCCAGGCAAG CCCAGCTGGA GTGGGATGAC TCCACCTCC CCTACTAGTG 2880
CCCAGGGGTC TGCTGCCTGG CCCACATGTC CCTTTTGTA ACCCTGACCC AGTGTATGCC 2940
CATGTCTATC ATACCTCACC TCTGATGTCT GTGACATGTC TGGGAAGGCC TTCTCCAGCT 3000
TCCTGGAGCC CACCTTTTAA GCCTTGGGCA CTCCCTGTGT TTCATCCATG GGGAAAGTCC 3060
AAGAAGCCCA GCATGGCCAT CAGTGAGGAC TTCAGGGTAG ACTTTGTCTT GTAGCCTCCA 3120
CTTCTGCCCT AAGTTCCCCA GCATCCTGAC TACCTGTCTG CAGAGTTTGC CTTTGTTTT 3180
TCCTGCAGGG AAGAAGGCC ACCTTTGTGT CACTCACCTC CCCAGGCTCA GAGTCCCCAA 3240
GGCCCTGGGG TTCCAACCTCA CTGTGCGTCT CCTCCACACA GACCAGTAGG TTCTCCTATG 3300
CTGACTCCAG GTTGCTTCAT ACAAGGAGGG TGGTTGAACT TCACACACGT AAGGTCTTAG 3360
TGCTTAACAG TTTAAAGGAA AGTCTTGTG GAGGCAGAAC TAAGTTTACA GGGAAAGGTA 3420
CACACATTCT CTCTCTCTCT CTCTCTCTGT CTATCTAGTT CCCAGCTTG GAGAGCCTTT 3480
CCCCTTGCTT CTTTCTGAGG CCATATAAGC TTATAAGAAA AGTCCCAAAC CAAGAATAGG 3540
TCCTTGGCCA CAAGCAGGGT CTGATCCCCC ATCAGAGCTA TCTGAGCCTG CCTGTCTGGG 3600
CACCTGTGTC AACCATGCAG CTACCTGCC AGGGGCACTC AGCAAACAGA ACCACAGGGC 3660
CCAGGAGGCA TTCCACACAG GCACTGCCCC AGGACAACAC AACAAGGACA GTCACAACAA 3720
GGACAACAAG GACACAACAC AACACACAAC AAGGACAGTC ACAACAAGCC TAGAGCCAGA 3780
AAGCAGATGG AAATGCTAAT GAGGTCAAAC GTAGGCTTCA TGGTGGGTGG AGTGGGGGTG 3840
GCTGGGCTCC CCCAGGACAG AGGGGACCTT GAGGTTGGCA AGGCTCTCAC CACTCAGCCT 3900
TATGTTCCCT TATCTCTAT CTTCCTCTT GAGAAAATAC ACGCTTTCTG CATGTATTAG 3960
AAACGCACGA GCTCCACCAA GTCTACAATG AAAGTTTGAA ATTAACTGC AAGGAATTAG 4020
AAGCATATTT GCAATCATTT CAGCTTCTTC TTCTTCTG TCATAAAGG AGGAACACTT 4080
TAGATAGAGG GCAATATAT CTGAAAACCT AATTTCTTTC TTTTCTTAT AAGGAAATCT 4140
TTTCCATCTC CATCTAACA TGCACAACCT GTGAAGAGAA TTGTTTCTAT AGTAACTGGT 4200
CTGTGATCTT TTGTGGCCAA GAGAATAGCA GGCAAGAATT AGGGCCTGA CAGAATTTCC 4260
ACGAAGCTCT GAGAACATGT TTGTTTCGAA TGTCTGATTC CTCTTTGTCA TCAATGTGTA 4320
TGCTCTGTCC CCATCCTTCA CTCCTCCTCA AGCTCACACC AATTGGTTTG GCACAGGCAC 4380
AGAGCTGGTC CCTAGTTAAG TGGCATTAT GTTAAAAAAA A

BFA1 Protein sequence (SEQ ID NO:15)

Gene name: calyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Protein
Accession #: NP_071414; Predicted Signal sequence: 1-20; Predicted TM domains: 832-848;
PFAM domains: cadherin domains: 48-151, 165-254; Summary: A type I membrane protein; a
member of the calyntenin family; is related to the FAT tumor suppressor; is likely an
adhesion molecule important in mammalian developmental processes and cell communication.

MLPGRLCWVP LLLALGVGSG SGGGGDSRQR RLLAAKVNKH KPWIETSYHG VITENNDTVI 60
LDPPLVALDK DAPVPFAGEI CAFKINGQEL PFEAVVLNKT SGEGRRLRAKS PIDCELQKEY 120
TFIIQAYDCG AGPHETAWKK SHKAVVHIQV KDVNEFAPTF KEPAYKAVVT EGKIYDSILQ 180
VEAIDEDCSP QYSQICNYEI VTTDVPFAID RGNIRNTEK LSYDKQHQQYE ILVTAYDCGQ 240
KPAAQDTLVQ VDVKPVCKPG WQDWTKRIEY QPGSGSMPLEF PSIHLETCDG AVSSLQIVTE 300
LQTNYYIGKC DRETYSEKSL QKLCGASSGI IDLLPSPSAA TNWTAGLLVD SSEMIFKFDG 360
RQGAIPDGI VPKNLTDQFT ITMWMKHGPS PGVRAEKETI LCNSDKTEMN RHHYALYVHN 420
CRLVFLLRKD FDQADTFRPA EFHWKLDQIC DKEWHYYVIN VEFPPVTLYM DGATYEPYLV 480
TNDWPIHPSH IAMQLTVGAC WQGGVTKPKQ FAQFFHGLSL SLTIRPGKME SQKVISCLQA 540
CKEGLDINSL ESLGQGIKYH FNPSQSILVM EGDDIGNINR ALQKVSYSINS RQFPPTAGVRR 600
LKVSSKVQCF GEDVCISIFE VDAYVMVLQA IEPRIITLRT DHFWRPAAQF ESARGVTLFP 660
DKIVSTFAK TEAPGDVKT DPKSEVLEEM LHNLDPCDIL VIGGDLDPKQ ECLELNHSEL 720
HQRHLDATNS TAGYSIYGVG SMSRYEQVLH HIRYRNWRPA SLEARRFRK CSELNGRYTS 780
NEFNLEVSIL HEDQVSDKEH VNHLLVQPPF LQSVHHPESR SSIQHSSVVP SIATVVIIS 840
VCMLVFVVM GYVRVRIAHQ HFIQETAAK ESEMDWDDSA LTITVNPMEK HEGPGHGEDE 900
TEGEEEEEEAE EEMSSSSGSD DSEEEEEEEG MGRGRHGQNG ARQAQLEWDD STLTPY

BFG7 DNA sequence (SEQ ID NO:16)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Nucleic Acid
Accession #: n/a; Coding sequence: <1-906 (stop codon underlined)

CGGGTCGACC CACGCGTCCG GGGAGAAAGG ATGGCCGGCC TGGCGGCGCG GTTGGTCCTG 60
CTAGCTGGGG CAGCGGCGCT GGCGAGCGGC TCCCAGGGCG ACCGTGAGCC GGTGTACCGC 120
GACTGCGTAC TGCACTGCGA AGAGCAGAAC TGCTCTGGGG GCGCTCTGAA TCACTTCCGC 180
TCCCGCCAGC CAATCTACAT GAGTCTAGCA GGCTGGACCT GTCGGGACGA CTGTAAGTAT 240
GAGTGTATGT GGGTCACCGT TGGGCTCTAC CTCCAGGAAG GTCACAAAGT GCCTCAGTTC 300
CATGGCAAGT GGCCCTTCTC CCGGTTCTCT TTCTTTCAAG AGCCGGCATC GGCCGTGGCC 360
TCGTTTCTCA ATGGCCTGGC CAGCCTGGTG ATGCTCTGCC GCTACCGCAC CTTCGTGCCA 420
GCCTCCTCCC CCATGTACCA CACCTGTGTG GCCTTCGCC TGGTGTCCCT CAATGCATGG 480
TTCTGGTCCA CAGTYTTCCA CACCAGGGAC ACTGACCTCA CAGAGAAAAT GGACTACTTC 540
TGTGCTTCCA CTGTATCCT ACACCTCAAT TACCTGTGCT GCGTCAGCCT CATCCGCTTC 600
GACTATGGCT ACAACCTGGT GGCCAACGTG GCTATTGGCC TGGTCAACGT GGTGTGGTGG 660
CTGGCCTGGT GCCTGTGGAA CCAGCGGCGG CTGCCTCAGC TGCGCAAGTG CGTGGTGGTG 720
GTCTTGCTGC TGCAGGGGCT GTCCCTGCTC GAGCTGTGTT ACTTCCACC GCTCTTCTGG 780
TCCCTGGATG CCCATGCCAT CTGGCACATC AGCACCATCC CTGTCCACGT CCTCTTTTTC 840
AGCTTTCTGG AAGATGACAG CCTGTACCTG CTGAAGGAAT CAGAGGACAA GTTCAAGCTG 900

GACTGAAGAC CTTGGAGCGA GTCTGCCCCA GTGGGGATCC TGCCCCCGCC CTGCTGGCCT 960
 CCTTCTCCCT CTCAACCCCT GAGATGATTT TCTCTTTTCA ACTTCTTGAA CTGGACATG 1020
 AAGGATGTGG GCCCAGAATC ATGTGGCCAG CCCACCCCT GTTGGCCCTC ACCAGCCTTG 1080
 GAGTCTGTTC TAGGGAAGGC CTCCCAGCAT CTGGGACTCG AGAGTGGGCA GCCCTCTAC 1140
 CTCCTGGAGC TGAACCTGGG TGGAACTGAG TGTGCTCTTA GCTCTACCGG GAGGACAGCT 1200
 GCCTGTTTCC TCCCCATCAG CCTCCTCCCC ACATCCCCAG CTGCCTGGCT GGGTCTGAA 1260
 GCCCTCTGTC TACCTGGGAG ACCAGGGACC ACAGGCCTTA GGGATACAGG GGGTCCCCTT 1320
 CTGTTACCAC CCCCACCCCT CCTCCAGGAC ACCACTAGGT GGTGCTGGAT GCTGTCTCTT 1380
 TGGCCAGCCA AGGTTACCGG CGATTCTCCC CATGGGATCT TGAGGGACCA AGCTGCTGGG 1440
 ATTGGGAAGG AGTTTCACCC TGACCRITGC CCTAGCCAGG TTCCAGGAG GCCTCACCAT 1500
 ACTCCCTTTC AGGGCCAGGG CTCCAGCAAG CCCAGGGCAA GGATCCTGTG CTGCTGTCTG 1560
 GTTGAGAGCC TGCCACCGTG TGTGCGGAGT GTGGGCCAGG CTGAGTGCAT AGGTGACAGG 1620
 GCCGTGAGCA TGGGCCTGGG TGTGTGTGAG CTCAGGCACT AGGTGCGCAG TGTGGAGACG 1680
 GGTGTTGTCT GGAAGAGGT GTGGCTTCAA AGTGTGTGTG GTGCAGGGGG TKGGTGTGTT 1740
 AAGCGTGGT TAGGGGAACG TGTGTGCGCG TGCTGGTGGG CATGTGAGAT GAGTGACTGC 1800
 CGGTGAATGT GTCCACAGTT GAGAGGTTGG AGCAGGATGA GGAATCCTG TCACCATCAA 1860
 TAATCACTTG TGGAGCGCCA CTGGGCCCAA GACGCCACCT GGGCGGACAG CAGGAGCTCT 1920
 CCATGGCCAG GCTGCCTGTG TGCATGTTCC CTGTCTGGTG CCCCTTTGCC CGCCTCTGTC 1980
 AAACCTCACA GGGTCCCCAC ACAACAGTGC CCTCCAGAAG CAGCCCCCTG GAGGCAGAGG 2040
 AAGGAAATG GGGATGGCTG GGGCTCTCTC CATCTCCTT TTCTCCTTGC CTTCGCATGG 2100
 CTGGCCTTCC CTCCAAAAC CTCCATTCCC CTGCTGCCAG CCCCTTTGCC ATAGCCTGAT 2160
 TTTGGGGAGG AGGAAGGGG GATTGAGGG AGAAGGGGAG AAAGCTTATG GCTGGGTCTG 2220
 GTTCTTCTCC TCCCAGAGG GTCTTACTGT TCCAGGGTGG CCCAGGGGCA GGCAGGGGCC 2280
 AACTATGCC TGCGCCCTGG TAAAGGTGAC CCCTGCCATT TACCAGCAGC CCTGGCATGT 2340
 TCCTGCCCCA CAGGAATAGA ATGGAGGGAG CTCCAGAAAC TTTCCATCCC AAAGGCAGTC 2400
 TCCGTGGTTG AAGCAGACTG GATTTTGTCT CTGCCCCTGA CCCCTTGTCC CTCTTTGAGG 2460
 GAGGGGAGCT ATGCTAGGAC TCCAACCTCA GGGACTCGGG TGGCCTGCGC TAGCTTCTTT 2520
 TGATACTGAA AACTTTTAAAG GTGGGAGGGT GGCAAGGGAT GTGCTTAATA AATCAATTCC 2580
 AAGCCTCAAA AAAAAAAAAA AAAAAAAAAA AAAAAA

BFG7 Protein sequence (SEQ ID NO:17)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 117-133, 179-195, 211-227, 235-251, 266-282, 296-312; PFAM domains: none; Summary: A type III membrane protein of unknown function; is adjacent to HER2 on the genome, and its overexpression in breast cancer is highly correlated with HER2 expression; may be used to predict HER2 overexpression and amplification.

RVDPRVRGER MAGLAARLVL LAGAAALASG SQGDREPYYR DCVLQCEEQN CSGGALNHFR 60
 SRQPIYMSLA GWTCDRDDCKY ECMWVTVGLY LQEGHKVPQF HGKWPFSRFL FFQEPASAVA 120
 SFLNGLASLV MLCRYRTFVP ASSPMYHTCV AFAWVSLNAW FWSTVFHTRD TDLTEKMDYF 180
 CASTVILHSI YLCCVRTVGL QHPAVVSAFR ALLLLMLTVH VSYLSLIRFD YGYNLVANVA 240
 IGLVNVVWWL AWCLWNQRRRL PHVRKCVVVV LLLQLGLSLE LLDFFPLFWV LDAHAIWHIS 300
 TIPVHVLFFS FLEDDSLYLL KESEDKFKLD

BCN4 DNA sequence (SEQ ID NO:18)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Nucleic Acid Accession #: n/a; Coding sequence: 143-874 (start and stop codons underlined)

GGGAGGGAGA GAGGCGCGCG GGTGAAAGGC GCATTGATGC AGCCTGCGGC GGCCTCGGAG 60
 CGCGGCGGAG CCAGACGCTG ACCACGTTCC TCTCCTCGGT CTCCTCCGCC TCCAGCTCCG 120
 CGCTGCCCCG CAGCCGGGAG CCATGCGACC CCAGGGCCCC GCCGCCTCCC CGCAGCGGCT 180
 CCGCGGCCTC CTGCTGCTCC TGCTGCTGCA GCTGCCCGCG CCGTCGAGCG CCTCTGAGAT 240
 CCCCAGGGG AAGCAAAAGG CGCAGCTCCG GCAGAGGGAG GTGGTGGACC TGTATAATGG 300
 AATGTGCTTA CAAGGGCCAG CAGGATGCC GGGAGCCCTG GGGCCAATGG 360
 CATTCCGGGT ACACCTGGGA TCCCAGGTCG GGATGGATTC AAAGGAGAAA AGGGGGAATG 420
 TCTGAGGGAA AGCTTTGAGG AGTCCTGGAC ACCCACTAC AAGCAGTGTT CATGGAGTTC 480
 ATTGAATTAT GGCATAGATC TTGGGAAAAT TGCAGAGTGT ACATTTACAA AGATGCGTTC 540
 AAATAGTGCT CTAAGAGTTT TGTTCAGTGG CTCACTTCGG CTAAAATGCA GAAATGCATG 600
 CTGTACGCGT TGGTATTTCA CATTCAATGG AGCTGAATGT TCAGGACCTC TTTCCATGTA 660
 AGCTATAATT TATTTGGACC AAGGAAGCCC TGAAATGAAT TCAACAATTA ATATTCATCG 720
 CACTTCTTCT TGGAAGGAC TTTGTGAAGG AATTGGTGTG GGATTAGTGG ATGTTGCTAT 780
 CTGGGTTGGC ACTTGTCAG ATTACCCAAA AGGAGATGCT TCTACTGGAT GGAATTCAGT 840
 TTCTCGCATC ATTATTGAAG AACTACCAAA ATAAATGCTT TAATTTTCTT TTGCTACCTC 900
 TTTTCTTATT ATGCCTTGA ATGGTTCACT TAAATGACAT TTAAATAAG TTTATGTATA 960
 CATCTGAATG AAAAGCAAAG CTAATATGT TTACAGACCA AAGTGTGATT TCACACTGTT 1020
 TTTAAATCTA GCATTATTCA TTTTGCTTCA ATCAAAAGTG GTTTCAATAT TTTTCTTAGT 1080
 TGGTTAGAAT ACTTTCTTCA TAGTCACATT CTCTCAACCT ATAATTTGGA ATATTGTTGT 1140
 GGTCTTTTGT TTTTCTCTT AGTATAGCAT TTTTAAAAA ATATAAAGC TACCAATCTT 1200
 TGTACAATTT GTAAATGTTA AGAATTTTCT TTATATCTGT TAAATAAAAA TTATTTCCAA 1260
 CAACCTTAAA AAAAAAAAAA AAAA

BCN4 Protein sequence (SEQ ID NO:19)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Protein Accession #: n/a; Predicted Signal sequence: 1-30; TM domains: none; PFAM domains: none; Summary: a secreted protein; has a mouse orthologue (see sequence below).

MRPQGPAASP	QRLRGLLLLL	LLQLPAPSSA	SEIPKGKQKA	QLRQREVVDL	YNGMCLQGPA	60
GVPGRDGSPG	ANGIPGTPGI	PGRDGFKEK	GECLRESFEE	SWTPNYKQCS	WSSLNYGIDL	120
GKIAECTFTK	MRSNSALRVL	FSGSLRLKCR	NACCQRWYFT	FNGAECGSL	PIEAIYLDQ	180
GSPMNSTIN	IHRTSSVEGL	CEGIGAGLVD	VAIWVGTCS	YPKGDASTGW	NSVSRIIEE	240

LPK

Mouse BCN4 Protein sequence (SEQ ID NO:20)
Gene name: ESTs; Unigene number: Mm.41556

XXXXAAPPQL	LLGLFLVLLL	LLQLSAPSSA	SEIPKVKQKA	LIRQREVVDL	YNGMCLQGPA	60
GVPGRDGSPG	ANGIPGTPGI	PCQDGFKEK	GECLRESFEE	SWTPNYKQCS	WSSLNYGIDL	120
GKIAECTFTK	MRSNSALRVL	FSGSLRLKCR	NACCQRWYFT	FNGAECGSL	PIEAIXXXXX	180
XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXSD	YPKGDASTGW	DSVSRIIEE	240

LPK